Sinophone Southeast Asia

Sinitic Voices across the Southern Seas

**Caroline CHIA and
Tom HOOGERVORST**

Sinophone Southeast Asia

Chinese Overseas

HISTORY, LITERATURE, AND SOCIETY

Chief Editor

WANG Gungwu

Subject Editors

Evelyn Hu-DeHart
David Der-wei WANG
WONG Siu-lun

VOLUME 20

The titles published in this series are listed at *brill.com/cho*

Sinophone Southeast Asia

Sinitic Voices across the Southern Seas

Edited by

Caroline Chia
Tom Hoogervorst

BRILL

*To Zacton Li, Celine Rebecca Hoogervorst,
and Ari Tea McFarland*

∴

Contents

Acknowledgements

Notes on the Language

In a book that aims to showcase and celebrate the full diversity of Sinitic and Sinitic-influenced languages in Southeast Asia, it would have been counterproductive for us to homogenize the spelling systems used in its chapters. Instead, we have relied on the judgment of our contributors. Some authors represent their data in standardized orthographies, such as Hànyǔ Pīnyīn 漢語拼音 for Mandarin, Jyutping 粵拼 for Cantonese, and Pėh-ōe-jī 白話字 for Hokkien. Others use an IPA-based representation – especially if no universally accepted spelling exists – or a common vernacular transcription. We follow the individual chapters, and their different conventions, in the concluding remarks. As a rule, the authors of this volume have respected the preferences of their fieldwork consultants and the spelling systems of their primary sources, especially if they contain important sociolinguistic information. As this book partly deals with historical varieties, we generally prefer traditional characters (正體字 *zhengti zi*, 繁體字 *fanti zi*) over simplified characters (简化字 *jianhua zi*). However, here too, we deviate from this principle if the consultants themselves prefer to use simplified characters. We represent non-standard characters exactly as we have encountered them in our source materials, refraining from attempts to "correct" them with the corresponding standard equivalents. In all cases, orthographic choices have been clearly indicated by the individual contributors.

Illustrations and Tables

Figures

Tables

Abbreviations

CAU	causative
CLF	classifier
EMH	Early Manila Hokkien
EXC	exclusive
EXP	experiential
F	female
G	generation
GPIRS	Gaginang Peng-im System
HT	historical works on Teochew
INC	inclusive
IPA	International Phonetic Alphabet
IT	Indonesian Teochew
JT	Jieyang Teochew
lit.	literally
M	male
NVOC	non-vocative particle
PART	particle
PL	plural
PN	personal name
PRC	People's Republic of China
POJ	Pėh-ōe-jī
POSS	possessive
PHH	Philippine Hybrid Hokkien
PRON	pronoun
RED	reduplication
ROC	Republic of China
REL	relative particle
SEQ	sequentiality marker
SG	singular
ST	Singapore Teochew
TCM	traditional Chinese medicine
TT	Thai Teochew
vol.	volume
VOC	Vereenigde Oostindische Compagnie (Dutch East India Company)

Notes on Contributors

Caroline Chia

Growing up in a family that speaks Hokkien and Teochew, Caroline's interest in *fangyan* sprouted from her maternal grandmother who came from Swatow. Grandma would listen to Teochew songs (*kêg⁴* 曲) on *Rediffusion* (a cable-transmitted radio station popular in Singapore in the mid-twentieth century, particularly for its *fangyan* programmes) and narrated her wartime stories to Caroline. Since her undergraduate days, Caroline began her research on traditional Chinese theatre. Starting with Hokkien-language theatre, this gradually expanded to Henghua, Teochew, and Hainanese. From being able to understand only a few lines to transcribing hours of performance, Caroline aims to learn more *fangyan* varieties to better document traditional Chinese theatre. Having spent more than a decade researching Hokkien theatre, her monograph *Hokkien Theatre across the Seas* was published in 2019.

Catherine Churchman

received her doctorate in the Research School of Pacific and Asian Studies at Australian National University and specializes in the premodern history of Tai peoples in the Sino-Vietnamese borderlands. She first learnt some Hokkien from a Malaysian roommate's friend in her first year as an undergraduate. After living in Taiwan, she took up the study of Taiwanese Hokkien, but became more attracted to Penang Hokkien in the mid-2000s as it was more widely spoken amongst people her own age. She began to listen to John Ong's weekly Penang Hokkien Podcast, asking Hokkien-speaking friends for help and advice as few resources for learning the language existed at the time. Her research on Penang Hokkien (specifically the mixed variety spoken by the Straits Chinese) is an offshoot of her long-term project of writing a Penang Hokkien-English dictionary. Catherine is currently lecturer in the Asian Studies Programme at Victoria University of Wellington, New Zealand.

Joanna Rose McFarland

is a PhD student at Nanyang Technological University in Singapore studying Teochew in Cambodia. Her research will provide a comprehensive reference grammar of Cambodian Teochew, while also examining the sociolinguistic situation of the group. Furthermore, she aims to compare the Cambodian variety to Teochew varieties spoken in Chaoshan and around Southeast Asia. Divergences will be analysed with respect to their relation to Khmer grammar,

with the purpose of determining what extent Khmer has influenced the grammar of Cambodian Teochew. Joanna graduated *cum laude* from the University of Washington in Seattle in 2010 with BAs in Linguistics and Communication. It was there where she met her husband, whose Teochew-speaking parents immigrated to the United States of America from Cambodia in the 1980s. Her new-found family sparked a keen interest in studying this largely unexplored language variety. She went on to complete her MA in Linguistics with Distinction in 2017 at the University of Hong Kong, where she provided a preliminary comparative analysis of Cambodian Teochew including evidence for contact-induced change.

Juliette Huber

has primarily focused on the languages of the Papuan Timor-Alor-Pantar (TAP) family spoken in the eastern part of Timor-Leste, in particular the closely related languages Makalero and Makasae. Since completing her PhD thesis, a descriptive grammar of Makalero, at Leiden University in 2011, she has studied among others landscape categorization and aspects of spatial language in the eastern Timor TAP languages and the diachrony of the TAP family. During fieldwork in Timor-Leste, she developed an interest for Timor Hakka when she made friends with the Li family of Lospalos, who regularly invited her round for dinners and whose hospitality helped her get through the occasional bout of fieldwork ennui. Hearing Hakka spoken in the family's home, Juliette was fascinated with how Timor-Leste's tiny Chinese minority had managed to maintain their language, and began wondering whether there was such a thing as a Timorese variety of Hakka. In the literature, she found nothing but occasional tantalizing hints, which led her to start collecting her own Timor Hakka data. To date, she is still very much in the beginning stages of her study, but is looking forward to devoting more time to this most neglected of Timor-Leste's languages. She is interested in whether a localized variety of Hakka has developed on Timor and what its characteristics are, and whether and to what extent Timor-Leste's variety differs from that spoken in Indonesia's West Timor.

Khin Khin Aye

is currently a research adjunct at the School of Design and Arts, Faculty of Business, Design and Arts of Swinburne University of Technology. Her areas of research interest include contact linguistics, education policy and nation building, learner interlanguage, and technology in language education. Her keen interest in Sinitic varieties derives from being a Fuzhou speaker married into a Hokkien family, whereas her love for Malay contact varieties and

the Sinitic influences they exhibit started with her MA research on the Malay of Myanmar speakers working in Brunei Darussalam. She researches Bazaar Malay and Baba Malay, investigating the extent of substratal influence from Hokkien on these varieties. She has presented papers on this topic at international conferences and published nine book chapters and two journal articles containing her research findings. She is also one of the contributors of *The Atlas of Pidgin and Creole Language Structures* by Oxford University Press.

Picus Sizhi Ding

With three academic degrees in Linguistics, Picus has devoted his research to languages of China, especially those less-studied. He is the author of three books, each dealing with aspect(s) of a specific language of China, including *A Grammar of Prinmi* and *Southern Min (Hokkien) as a Migrating Language*. Born in a Hokkien family in Myanmar and grown up in Macao, he has found himself in a multilingual setting since his early childhood. It is precisely this unusual experience that motivated him to pursue studies in Linguistics throughout his tertiary education. While Hokkien is the language he first acquired and spoke, it is not his dominant language, since the use of Hokkien has largely been confined to the domain of family. After decades of conducting fieldwork on minority languages of Yunnan, he observed a striking parallelism in decline of linguistic vitality between indigenous languages of ethnic minorities and heritage languages of migrant communities. This prompted him to investigate the current sociolinguistic status of Hokkien in selected parts of Southeast Asia, Taiwan, and Fujian, based on his personal experiences in these Hokkien-speaking regions. The result of this study is published in *Southern Min (Hokkien) as a Migrating Language: A Comparative Study of Language Shift and Maintenance across National Borders*.

Tom Hoogervorst

has worked primarily on Malay and Javanese, including the varieties of these languages spoken by Indonesia's Chinese communities in past and present. His latest book, *Language Ungoverned: Indonesia's Chinese Print Entrepreneurs, 1911–1949*, focuses on the Malay print culture of late-colonial Indonesia, which was dominated by ethnic Chinese and featured a hybrid, Hokkienized type of Malay. As a scholar of language contact and lexical borrowing, Tom is also an armchair linguist when it comes to the languages historically in contact with those of Indonesia, including Hindustani, Tamil, Persian, Arabic, and Sinitic varieties. Given the prominence of Fujianese settlers throughout Indonesia's history, his research pays particular attention

to the way Hokkien and Malay have influenced each other. His relatively recent interest in Mandarin stems from the observation that this language differs in interesting, historically revealing, and largely understudied ways in mainland China, Taiwan, Singapore, Malaysia, Indonesia, and other parts of Southeast Asia.

Introduction

In August 2020, *The Royal Singapore* – a news platform "that celebrates the lives of interesting and inspiring Singaporeans and those who live in and contribute to Singapore" – shared a video of the local polyglot Mdm. Rasamal, better known as Aunty Rose.[1] In slightly over eight minutes, one can see her switch seamlessly between Cantonese, Hokkien, Teochew, Bazaar Malay, Mandarin, and Tamil. As several netizens pointed out, this brisk, cheerful octogenarian symbolizes the "kampong spirit" of early Singapore: hybrid, plurilingual, resilient, kind, and tolerant to difference. Singaporean commenters were quick to point out that her Cantonese – or, for that matter, any of the other languages she spoke – was better than theirs, even though she was of mixed parentage and thus not "officially" Chinese in the nation's dominant framework on race. Many were reminded of their own grandparents and expressed regret that such enviable levels of proficiency have decreased in recent times. But everyone agreed that the video was quintessentially Singaporean (and Malaysian, as some added). The Cantonese of Aunty Rose features delightful local expressions like *hung mou meng* 紅毛名 'Western name (as opposed to Chinese name)', *yoeng me si* 羊咩屎 'goat's droppings (the name of a traditional snack)', and *tong yan* 唐人 'ethnic Chinese people'. It also contains the English words *auntie*, *card*, *happy*, and *RC* (resident's committee), alongside several phonologically integrated loans from Malay: *aa daap* 'nipa palm leaves' (*atap*), *baa saa* 'market' (*pasar*), *daa bi* 'but' (*tapi*), *gam bong* 'traditional village' (*kampung*), *lo di* 'bread' (*roti*), and *o dang* 'to owe money' (*utang*).

Not so long ago, linguistic repertoires of such magnitude were the norm for many Southeast Asians – especially in families with ancestors from overseas – and for some they still are. This book highlights the language practices of Southeast Asia's diverse Sinophone communities, paying attention to typological characteristics, sociolinguistic histories, and correlations with culture and identity. We aim to contribute, at once, to the scholarly literature on Chinese languages outside China and to the field of Southeast Asian linguistics, in which Sinitic varieties have received considerably less attention than languages deemed "indigenous" to the region. We adopt the term "Sinitic" (華 *hua*) for Chinese in a broad ethno-linguistic sense, and prefer the term regional varieties (方言 *fangyan*) to "dialects". The Sinitic languages, language histories, and sociolinguistic practices of Southeast Asia have yet to be analysed

---

1 https://www.facebook.com/TheRoyalSingapore/ (accessed 5 August 2020). See also https://youtu.be/m-SK9rLYnCs (accessed 6 May 2021).

comparatively.[2] Many of Southeast Asia's unique Sinitic languages are now endangered, as fewer and fewer individuals retain fluency in the tongues of their grandparents, particularly amidst the hegemony of national languages, Mandarin, and English. Yet they constitute important linguistic heritage, are closely intertwined with (often equally endangered) localized cultural practices, and crucial to the region's grassroots histories.

The diverse stories of Southeast Asia's Chinese communities cannot be fully grasped by prioritizing their politics and socioeconomic trajectories over language and culture, nor through well-established yet ultimately reductive tropes such as "overseas Chinese" or the "Chinese diaspora".[3] We believe the term "overseas Chinese" is to be avoided in reference to contemporary times, as it connotes people who reside overseas but are still intimately connected to China in terms of their political allegiance. Most Chinese-descended communities in Southeast Asia, by contrast, associate themselves predominantly with their country of residence. The expression "Chinese diaspora" suffers from similar limitations, but is also cryptic in a Southeast Asian context, where many ethnic Chinese have formed new diasporas – for example to Australia, Europe, or within Southeast Asia – in the face of precarity under hypernationalist regimes.

This book aims to contextualize the plethora of Sinitic linguistic practices and expressions – or "voices", as we have come to call them – in Southeast Asia by bringing together perspectives and empirical data from scholars of various geographical and disciplinary backgrounds. Each chapter approaches language in conversation with history and identity. This is especially useful in minority contexts, where the three components are often seen as interconnected nodes within the framework of heritage. Examining language, history, and identity together allows us to jump between different scales of perspective, linking specific localized idioms and language practices of certain families (or even individuals) to global patterns of Chinese migration, national Southeast Asian language policies, and enduring legacies of interethnic contact.

---

2   The only comparative study known to us is a descriptive, data-rich volume edited by Li Rulong (2000). For an insightful edited volume on multilingualism in Chinese-descended communities worldwide, see Li Wei (2015).

3   See Shih (2013) for a detailed argument against the usage of these terms.

1   Broadening the Sinophone

In the vibrant field of Sinophone Studies, which examines the cultural productions of Sinophone and Sinicized (漢化 *hanhua*) communities,[4] Southeast Asia has by no means been ignored. Many of its proponents have advocated for a departure from what might be called mainland-centrism and an appreciation of the great internal diversity of those labelled as ethnic Chinese, in Southeast Asia and elsewhere.[5] E.K. Tan's *Rethinking Chineseness: Translational Sinophone Identities in the Nanyang Literary World* (2013) is a pioneering study on Chinese-language writers based in Borneo, Malaysia, and Singapore that reflects on the notion of "Chineseness", identity, and the evolution of local cultures. Alison Groppe's *Sinophone Malaysian Literature: Not Made in China* (2013) highlights literary expressions on being Chinese forged in Malaysia. Brian Bernards' *Writing the South Seas: Imagining the Nanyang in Chinese and Southeast Asian Postcolonial Literature* (2015) examines the way Southeast Asia is described and imagined in China and Sinophone Southeast Asia. Chia-rong Wu's *Supernatural Sinophone Taiwan and Beyond* (2016) studies comparatively the Sinophone literature of Taiwan and Malaysia, providing insights into the dynamics of storytelling and religious beliefs such as the deification of ghosts. Hee Wai-Siam's *Remapping the Sinophone: The Cultural Production of Chinese-Language Cinema in Singapore and Malaya before and during the Cold War* (2019) explores Chinese-language cinema in Singapore and Malaya. Such studies illustrate the complexities of literature, film, and to some extent the interaction of Sinitic and non-Sinitic languages, although they have remained somewhat silent on Sinitic music, performed arts, and "folk" (民間 *minjian*) texts.[6]

This book adds language itself to this expanding range of topics. Doing so, especially with a focus on Southeast Asia, underscores the benefits, limitations, and future potential of the Sinophone as a concept-in-progress. Some may find that the term has become overtheorized, drifting away from other "-phones" – such as the Francophone or Lusophone – and prioritizing postcolonial literature and liminal cultural productions. To interpret what is found

---

4   Shih (2004, 2007), Tsu & Wang (2010), Shih et al. (2013).

5   This enterprise yielded the Cambria Sinophone World Series edited by Victor H. Mair, which foregrounds "Sinitic-language cultures and communities born of colonial and postcolonial histories that lie on the margins of geopolitical nation-states across the world". http://www.cambriapress.com/cambriaseries.cfm?template=85 (accessed August 2020).

6   In contrast to the literary texts covered in Sinophone Studies, these folk texts – particularly those pertaining to still vibrant religious customs, traditions, and theatre – play a key role in Southeast Asia's Sinophone practices.

in Southeast Asia, we assert that linguistically encoded processes of hybridity and localization are crucial. At the same time, language-centric approaches add new layers of complexity. If one views language as the decisive factor in defining and demarcating the Sinophone, one would have to exclude the majority of cultural productions – including those dealing with Chineseness – by ethnic Chinese in Indonesia, as they are written in non-Sinitic languages.[7] The position of mixed languages such as Baba Malay and Bazaar Malay (Aye, this volume) or plurilingual publications (Hoogervorst, this volume) would be ambiguous. Yet the Sinophone has also been characterized as the product of a "condition of exile, diaspora, minoritization, and hybridity that resists incorporation both into China and into the place of residence" (Shih 2004: 26). In that case, many modern Chinese-language productions from Southeast Asia are not Sinophone either, as their authors would actually welcome a long-awaited incorporation into the nation-state, while the PRC – at least in the 1950s and 1960s – was too indifferent about them to actively pursue incorporation into its literary canon.[8]

Yet more than delineating who or what constitutes the Sinophone, the chief task at hand is to explore the conceptual value of "Sinophone Southeast Asia". A major epistemic advantage of the Sinophone, as we see it, is its potential to encompass both written and oral language practices, together with its openness to non-standard languages and regional variability.[9] In this regard, Southeast Asia's multifaceted storehouse of experiences and case studies – centring on transregional circulations, layered histories of mobility, and exceptional plurilingualism – invites a move beyond the common dichotomy of mainland China versus "the diaspora". The voices we have gathered under the rubric of Sinophone Southeast Asia encompass cultural productions in various Sinitic and Sinitic-influenced languages as well as the underlying practices of language contact and plurilingualism. This fruitful combination unlocks new avenues to investigate not only textuality, but also the materiality of language, including the various ways in which Southeast Asia's Sinitic languages can be written down or used in performances, and the different sociolinguistic implications of these choices. On a conceptual level, Sinophone Southeast Asia invites us to think more radically about geographies, routes, contacts, boundaries, and identities as embodied by the region's Sinitic and Sinitic-influenced languages.

---

7   See Chandra (2015) on this point.

8   See Stenberg (2017) on this point in relation to Chinese-language literature in Indonesia.

9   This is amply demonstrated in the work of Tsu (2010). We must also call attention to the work of Leow (2016) and Tam (2020) for a radical reappraisal on the role of regional Sinitic varieties in, respectively, Malaysia and mainland China.

The above example of Aunty Rose and her extremely rich Sinophone repertoire provides a case in point. It reminds us that the Sinophone is not exclusively defined or produced by people categorized as Han (cf. Shih 2013: 7). If anything, it inherently extends to other ethnic groups, including non-Chinese, as various chapters of this volume instantiate.

We also propose that Sinitic varieties other than Mandarin merit closer attention in Sinophone Studies. Mandarin – also known as Huayu 華語, Guoyu 國語, or Putonghua 普通话 – has become the default lingua franca for Chinese communities inside and outside China. For the latter group, it may also serve as a reminder of their ethnicity, or the fact that their ancestors came from China. Regional languages or *fangyan*, by contrast, point specifically to one's ancestral province, county, or even village. They help speakers – and, to some extent, rememberers – relate to their local identity on a more intimate level. Regional languages also enable Chinese-descended Southeast Asians to experience and understand the cultural practices and customs that have been passed down to them intergenerationally. We must emphasize here that Sinophone Southeast Asia differs from other parts of the world in its remarkable linguistic diversity. Unlike early Chinese migration to Australia and the Americas, which was dominated by Cantonese communities, Southeast Asia exhibits a complex makeup involving speakers of Hokkien, Teochew, Cantonese, Hainanese, Hakka, Henghua 興化, Hokchia 福清,[10] and many more.

2   Lessons from Southeast Asia

This is also a book about Southeast Asia. Whereas some volumes dealing with this part of the world start by problematizing the region, we simply remind our readers that this geographical unit made perfect sense to the ancient Chinese, who designated it as the "Southern Seas" (南洋 *Nanyang*). Sojourning or settling in the Nanyang – an activity once designated as "travelling overseas to [the land of the] southern barbarians" (*guofan* 過番) – provided opportunities not found in China. While a considerable part of these travellers never returned, China was frequently imagined, by them and their descendants, as the semi-mythical homeland, especially before the emergence in Southeast Asia of modern nation-states. More importantly for the purposes of this book, the Nanyang made regional Sinitic varieties mobile and eventually birthed mixed languages and cultures unique to what is now known as Southeast Asia.

---

10   Known in Singapore as *Hokchew*, this is a subvariety of Fuzhounese 福州話 (*Fuzhouhua*).

Inspired by Brian Bernards' concepts of "archipelagic imagination" and "continental imagination" (2015), we look at Southeast Asia as a maritime gateway, connecting with China but also facilitating regional networks that often transcended colonial and later national boundaries. By thinking "across the Southern Seas", we aim to look beyond official languages and national cultures, focusing instead on marginalized, creolized expressions of language and culture. For this reason, Chinese-descended Southeast Asians are not the only actors this volume takes interest in. In Chapter 2 on Timor Hakka, we are reminded that Creole Portuguese once served as a link-language between Timor and Macao. Meanwhile, Chapter 4 recalls that Bazaar Malay was the unofficial lingua franca of pre-independent Singapore, used by people of various backgrounds. Despite their marginalized status at present, these two non-Sinitic languages were once responsible for the interaction between Southeast Asia's innumerable ethnicities. They are relevant to the project of Sinophone Studies by virtue of their trans-ethnic, multicultural, and creolized characteristics, as well as their cross-pollination with the region's Sinitic varieties.

In addition to internal diversity, we believe that Chinese communities in Southeast Asia display a number of characteristics not seen to the same extent in North America, the Caribbean, Peru, Mauritius, South Africa, India, and other parts of the world. Southeast Asia and China have been connected since antiquity and increasingly so from colonial times, yielding processes of economic, cultural, and linguistic convergence of unparalleled time depth.[11] Unsurprisingly, Southeast Asia's Chinese communities have preserved various traditions forgotten and sometimes violently supressed in the PRC, especially during the Cultural Revolution (1966–76). They developed an even greater number of uniquely local ones, and at times spearheaded the introduction of foreign ideas and commodities into East Asia (Ding, this volume). Since overseas travel was historically a male prerogative, marriages between Chinese men and local women were common across Southeast Asia. These long legacies of admixture make it impossible to determine the precise number of ethnic Chinese in the region, but as a whole, Southeast Asia is certainly home to the largest Chinese-descended population anywhere in the world outside the PRC.

Across the Nanyang, Chinese migration consisted of different waves, with traders and sojourners (from medieval times), tax farmers, low-wage and

---

11   See Alves (2021) for an overview of linguistic connections between Southeast Asia and China. A similar argument could be made for Japan, Korea, Mongolia, Vietnam, and other countries in direct proximity to China, yet in these instances Sinicization was also a religious, administrative, and/or expansionist project.

often indentured labourers (from the mid-nineteenth century), and crafts-men constituting some of the major demographics. Sinitic communities that might not have interacted much in China often lived side by side in Southeast Asia, even though most professions were monopolized by specific ethno-linguistic groups. Chinese urban groups typically had their own quarters, temples, cemeteries, schools, meeting halls (會館 *huiguan*), clan associations, and leaders (in colonial times often referred to as "captains"). Integration into local populations generally took place more rapidly in the countryside. While cultural exchange and friendly relations with non-Chinese groups were com-mon throughout Southeast Asia, Chinese people have often been targeted for racialized violence in more tumultuous times. Colonial authorities regularly dismissed them as untrustworthy middlemen and exploited them through all sorts of ethnic quotas. After independence, several Southeast Asian regimes struggled to incorporate the ethnic Chinese minority into the nation-state. Mobilizing racialized mechanisms of exclusion – of the type that had infu-riated the region's first generation of anti-colonialists – policies to outlaw Chinese schools, newspapers, and organizations were implemented in sev-eral parts of postcolonial Southeast Asia and led to a dramatic erasure of local Sinitic languages and cultures. The economic and geopolitical "Rise of China" – and previously the economic prowess of Taiwan – heralded reconfigurations of Chineseness over the past decades. Yet these ongoing processes of "resinicization" have done little to arrest the marginalization of Southeast Asia's hybrid Sinitic languages and localized cultural practices, and may in fact have contributed to it. Other developments, such as the PRC's enormous infrastructural investments, expansionism in the South China Sea, and – especially in Cambodia and Myanmar – far-reaching political interfer-ence, have added to the complexity of contemporary (Sino-)Southeast Asian attitudes towards mainland China.

While Mandarin has been promoted and taught in Southeast Asia from the early twentieth century, and Yunnanese speakers of Southwestern Mandarin – a variety quite remote from what would later become the standard – arrived in the nineteenth century in parts of northern Thailand, Myanmar, and Laos, it is a relatively recent arrival to the region's linguistic landscape (Wang 2012; Sai 2016). Maritime connections, which have long linked coastal Southeast Asia with China's southern provinces, resulted in the large-scale immigration of Southern Min, Cantonese, and Hakka speech communities from early-modern into late-colonial times. In many cases, the survival of these Sinitic languages and cultural practices was ensured by new immigrants. There were also strong founder effects, in which newcomers had to adapt to existing norms and prac-tices. Varieties of Southern Min gained a dominant status in much of Southeast

Asia, in particular those belonging to the "Hokkien" 福建[12] and "Teochew" 潮州 groups. But the situation soon became more complicated, as descendants of Chinese migrants gained greater fluency in Southeast Asia's local languages. Even after adopting these new languages, however, Chinese families typically retained the cultural, culinary, and kinship terms from their ancestral Sinitic variety. The resultant manifestations of lexical borrowing form a recurring theme throughout the present volume.

3   Sinitic in All Its Diversity

As mentioned previously, we use the term "Sinitic" to refer to languages conventionally labelled with the more politically charged term "Chinese".[13] Most Sinitic languages discussed in this volume are the result of two variables: their specific regional origins in China and their history of language contact after migration. In addition to the overemphasis of Mandarin at the expense of *fangyan*, it must be kept in mind that Southeast Asia often exhibits specific subdialects rather than generic forms of "Hokkien", "Teochew", and "Hakka", as will be pointed out in the individual chapters. These varieties subsequently underwent phonological and/or lexical influence from languages in contact, such as Malay, Thai, Burmese, Khmer, or Vietnamese, but also other Sinitic varieties. Needless to say, they also donated numerous loanwords to the non-Sinitic languages of the region. As a result, Sinitic varieties in Southeast Asia differ substantially from their counterparts in China, but also from each other. After several generations of language contact, lexical, phonological, and grammatical interference from the surrounding languages often makes it difficult for Southeast Asia's Sinophone communities to understand each other even when speaking the same "dialect" (although such claims of limited comprehension also partly reflect social constructability). In addition to regional differences, many Sinitic languages – in Southeast Asia and elsewhere – distinguish between colloquial and literary registers, especially in performance genres. At the same time, increased contacts with mainland China, Hong Kong,

---

12   We use this term in its "exonymous" Southeast Asian sense. In linguistics, the Hokkien dialect continuum is known as the Quanzhang 泉漳 sub-division of the Southern Min (Fujian) varieties. To Hokkien speakers, the language might be designated as *Hok-ló-ōe* 福佬話 'Hoklo language', *Bân-lâm-gú* 闽南语 'Southern Min', or *Tâi-gí* 臺語 'Taiwanese', whereas *Hok-kiàn* 福建 itself refers to the broader Fujian region and its many other languages.

13   A similar view is provided by Ng (2013: 89): "since the Chineseness of this language should not be emphasized, we may as well call it the 'Sinitic language'".

and Taiwan have created a situation of diglossia, in which generic East Asian varieties enjoy a higher status than localized Southeast Asian ones.

Word histories constitute another unifying thread across this book. Several categories come to the fore. Across Southeast Asia, the terms used for Chinese, indigenous, and other communities differ from their PRC counterparts. Aunty Rose's usage of *tong yan* 唐人 evokes the Tang Dynasty, with which people from China's southern provinces strongly identified.[14] In a Southeast Asian (and Taiwanese) context, the term *Zhongguo ren* 中國人 'Chinese person' denotes PRC nationals rather than local Chinese, who may instead be designated as *Huaren* 華人. Since not all parts of Southeast Asia have been equally connected to China, a number of modern concepts exhibit distinct words in the Sinitic varieties of Southeast Asia, although closer contacts with China are now erasing this layer of unique, locally coined vocabulary (Churchman, this volume). Another clear point of divergence is the nomenclature for fruits, vegetables, and other locally specific products and commodities (see the chapters of McFarland and Churchman). Across the Sinitic varieties of Southeast Asia, these are often designated by loanwords or loan translations. If they are at all written down, the chosen characters might reflect a degree of phono-semantic matching. Consider, for example, Hokkien *kam-á-bit* 柑仔蜜 'tomato' (lit. 'tangerine honey') and *âng-mô·-tan* 紅毛丹 'rambutan' (lit. 'red-haired crimson') – borrowed respectively from Tagalog *kamatis* and Malay *rambutan* – which have also spread to non-Southeast Asian varieties of the language (and in the case of *rambutan* to Mandarin). The element *âng-mô* 紅毛 'red-haired' has itself adopted a wide range of meanings related to light-skinned Europeans, as discussed in detail in the chapters of Ding and Churchman. The aforementioned use of Singaporean Cantonese *hung mou meng* 紅毛名 'Western name' is but one of many incarnations of this versatile term, which has also entered Singaporean English (*angmoh*). Another well-known example among Southeast Asia's Sinitic varieties is 巴刹 (*pa sat, baa saa, ba sha*, etc.) 'market', from Malay *pasar*. In addition, we encounter hybrid expressions, consisting of one Sinitic and one "indigenous" Southeast Asian element (see Aye, this volume).

Historically, Chinese-descended Southeast Asians had relatively low levels of literacy and most localized Sinitic tongues were confined to the oral domain. This has long been the case even in China's southern provinces, whose *fangyan*

---

14   We may recall in this regard that China's southern provinces were frontier regions prior to their inclusion into the Tang cultural sphere. This is also exemplified by the related term "Tang Mountain" 唐山 for mainland China, which can be found in multiple Southeast Asian settings.

were not historically transformed – if not tamed – into written languages. For this reason, it makes sense to decentre the written word in order to gain a fuller grasp of the Nanyang's diverse Sinophone voices. That being said, we do not advocate for a complete move away from writing, as we believe the research community should pay attention to orality, textuality, and the interplay between them. In Hokkien varieties in particular, the co-existence of literary and colloquial readings for large numbers of words underline the importance of Chinese characters. In many cases, it was through the written word rather than oral transmission that Southeast Asian concepts found their way into the broader Sinosphere (Ding, this volume). As writing was not standardized, we find several non-mainstream characters in locally-authored sources. Some common examples of unique characters in Southeast Asian Hokkien include *ngé* 嬈 – 'concubine' in Java (from Malay *nyai*) and *leng* 朥 'rubber' in Malaya. Vernacular words were often written down in a variety of ways. We can encounter multiple choices of characters, for instance, of Zhangzhou Hokkien *ka-choáh* 'cockroach' and *ńg-kong* 'grandfather'.[15] For this reason, we have decided to also pay attention to the "heterogeneity of Sinophone writing" (Ng 2013: 76).

One topic that has received little attention thus far is the difficult enterprise of reconstructing Chinese words on the basis of romanized data, for example in texts by local Southeast Asians or foreign observers dealing with Chinese communities, or by Chinese authors using non-Sinographic scripts. Some of the manuscripts of Hokkien-language opera in the Philippines are a case in point (Chia, this volume). To accurately interpret such romanized data requires fluency in different registers and a deep understanding of the context in which they were produced. The same is true for occasional passages of Hokkien or Hakka in the Malay of Chinese-Indonesian authors, especially during the 1920s and 1930s. Lacking Chinese characters, such phrases can be quite opaque even for speakers of those varieties, as they occur in an unfamiliar romanization and lack tone marks, so that almost every word comes with several possible interpretations. These difficulties have led most scholars to ignore such texts, even though they were obviously understood at the time they were published. Arguably, they embody the Sinophone in its utmost diversity.

Of course this book has its omissions, which include relatively small Sinitic communities such as the Hainanese, Henghua, Hokchia, and Southeast Asian Chinese who have migrated to China or Vietnam. Five of the volume's chapters

---

15   *Ka-choáh* 'cockroach', for example, could be written as 虼蚻, 甲由, 蟉蟊, 蟉蟻, and
        狡蟊; and *ńg-kong* 'grandfather' as 阿公, 俺公, 映公, 安公, and 翁公. The pronuncia-
        tions *íng-kong* and *áng-kong* are also attested.

touch upon different Hokkien varieties. This reflects the historical importance of this language, but also academic priorities within Sinophone Southeast Asia. We believe that other regional languages and cultures are equally significant and hope that more data will be collected in future research. The written traditions of non-Sinitic languages in Sinographic scripts – such as Vietnamese and Zhuang – deserve a comparative study of their own and are not discussed here. We expect that continued attention to this field will fill some of these gaps in the future. Such an endeavour requires a fair deal of urgency, as many of the Sinitic varieties studied in this volume are on the brink of extinction and are unlikely to be passed on to the next generation. Even within the respective speech communities, some people deem the language of their parents and grandparents "impure" and best replaced by a standardized variety. At the same time, we were impressed by the levels of determination we saw among the various Sinitic communities in Southeast Asia to preserve their languages against the odds. In fact, in almost all cases, the communities themselves – both in Southeast Asia and in the "diaspora" worldwide – have encouraged and invited us to document their unique varieties, compare them with those of others, and detect regional patterns. They also made it clear that language, history, and identity cannot be separated if the academic output is to be meaningful for them.

Most chapters in this volume, hence, describe a poorly known local history, followed by data from an understudied linguistic variety, and sociolinguistic attitudes inside and outside the community in question. They also bring to the fore some of the agents shaping language history on a micro level, including translators, playwrights, authors of linguistic material, and fieldwork consultants. Though many of the volume's chapters present ongoing and early-stage research, we feel their publication is of a highly timely nature. We hope to inspire researchers and communities to carry out additional research, encourage comparisons across Southeast Asia, and work with elderly speakers before their knowledge can no longer be passed on. The individual contributors have made an effort to explain their methods, reflect on them, and convince the interlocutors within the community that their linguistic varieties are worthy of structural attention. In view of this scholarly-community collaboration, we collectively found it important that this volume was made publicly accessible.

4   Chapter Outline

Chapter 1 by Picus Ding delves into the historical context of Sinophone Southeast Asia, demonstrating how etymologies can illuminate the complex

trajectories – sometimes dating back to the fifteenth century – of a number of common words. It traces some of the earliest documented interactions between East Asian, Southeast Asian, and European communities and the resultant linguistic cross-pollination. While the economic contributions of Southern Min tradesmen are well understood – in Southeast Asia, the term "Hokkien" has ubiquitously come to refer to people from southern Fujian if not the whole of Fujian – their long-lasting sociolinguistic and cultural impact is not to be ignored. The two best known examples discussed in this chapter are the words "angmoh" and "pidgin", which have a history of several centuries and were probably coined by Southeast Asia-based Chinese. Ding makes the original claim that the term "pidgin" is influenced by Southern Min.

Chapter 2 by Juliette Huber explores a lesser-known group: the Hakka 客家 or 'guest people' in East Timor (Timor-Leste). This now independent state – situated on the periphery of the Nanyang – has received little international attention beyond the political upheavals leading up to its independence at the turn of the twenty-first century. Timor Hakka is not recognized among the state's official languages and is largely confined to the domestic sphere. In her pioneering description of the characteristics of Timor Hakka, Huber makes comparisons with Hakka varieties from other areas, including Meixian, Hong Kong, and Malaysia. Like many endangered Sinitic languages, the unique Hakka variety of East Timor constitutes a significant marker of the group's identity, including for those who live abroad.

Joanna McFarland's study of Cambodian Teochew, Chapter 3, fills an equally important gap. This sparsely researched variety is spoken by a "majority within a minority": the Teochews constitute the largest group within Cambodia's Chinese population but nevertheless remain a minority in the country. McFarland studies the linguistic features of Cambodian Teochew and examines its interaction with Khmer, the official language of Cambodia. The influence of Teochew on Khmer is a testimony to the historical importance of this community and their now endangered language. The chapter's attention to Cambodian Teochew speakers of different age groups and sites of residence (within and outside Cambodia) provides important insights into the sociolinguistic differences of this community.

Chapter 4 by Khin Khin Aye explores Baba Malay and Bazaar Malay in Singapore, illustrating an important episode in the island's linguistic history, especially before English became the lingua franca. Aye's study foregrounds Hokkien as the major substrate language of Baba Malay and Bazaar Malay, as can be seen, among others, in the arenas of business and kinship. This chapter

underlines the importance of substrate languages, which arguably deserve more attention in Pidgin and Creole Studies, by calling attention to the historical predominance and ubiquity of Hokkien speakers in Singapore and their impact on its creolized Malay vernaculars.

The next chapters highlight additional facets of Southeast Asia's rich Hokkien legacy. Chapter 5 by Catherine Churchman describes a variety known as Penang Hokkien or Northern Malaysia Hokkien. This language exhibits influence from Malay, English, and other languages. In addition, it has coined several localized Hokkien terms. An examination of its vocabulary yields fascinating insights into the history of Penang and Malaysia in general. The linguistic versatility of its speakers is ongoing. Whereas older speakers use numerous terms unique to Penang, the vocabulary of younger speakers reflects influence from Sinitic varieties outside Malaysia. As such, the chapter's examination of Penang Hokkien adds ample substance to the observation that the Sinophone "is a place-based, local culture, in dialogue with other cultures of that location" (Shih 2013: 8).

Chapter 6 by Caroline Chia traces the development of *Kaoka* 高甲 or *Gaojia* opera in the Philippines. This theatrical form originates from southern Fujian and has spread to Southeast Asia during the late nineteenth century. Once a popular form of entertainment for the Chinese communities across the Nanyang, the Philippines is the only locale in which Kaoka is still performed to this day. The Kaoka playscripts, written in romanized Hokkien, offer unique and sparsely researched data on localized Hokkien. These sources exemplify how the Sinophone can be produced by different communities, including non-Chinese people. Their script provides rare insights into non-standard Sinitic writing practices in Southeast Asia.

Chapter 7 by Tom Hoogervorst unearths an early resinicization discourse. It compares a number of Mandarin textbooks published in Java during the first half of the twentieth century. Written with the aim to reconnect Indonesia's Chinese communities with their perceived ancestral heritage, these books feature surprisingly little of the standard Mandarin that we know today. Instead, they showcase the competing types of Mandarin historically taught in Indonesia. It is of additional interest – especially with regard to the plurilingual focus of Sinophone Studies – that the language of instruction was Malay. On the surface, the Hokkienized colloquial Malay in which these textbooks were written resembles Baba Malay as discussed by Aye in Chapter 4. Nevertheless, both the Malay and the Hokkien of Java display various local characteristics not found elsewhere in Southeast Asia.

References

Alves, Mark J., 'Linguistic influence of Chinese in Southeast Asia', in Paul Sidwell and Mathias Jenny (eds.), *The Languages and Linguistics of Mainland Southeast Asia: A Comprehensive Guide* (Berlin: De Gruyter Mouton, 2021): 649–71.

Bernards, Brian, *Writing the South Seas: Imagining the Nanyang in Chinese and Southeast Asian Postcolonial Literature* (Seattle: University of Washington Press, 2015).

Chandra, Elizabeth, 'Review of *Rethinking Chineseness: Translational Sinophone Identities in the Nanyang Literary World*, by E.K. Tan. Cambria Press, Amherst, NY, 2013', *South East Asia Research*, 23/2 (2015), 263–65.

Hee Wai-Siam, *Remapping the Sinophone: The Cultural Production of Chinese-Language Cinema in Singapore and Malaya before and during the Cold War* (Hong Kong: Hong Kong University Press, 2019).

Leow, Rachel, *Taming Babel: Language in the Making of Malaysia* (Cambridge: Cambridge University Press, 2016).

Li Rulong 李如龙 (ed.), 东南亚华人语言研究. (Beijing: 北京语言文化大学出版社, 2000).

Li Wei (ed.), *Multilingualism in the Chinese Diaspora Worldwide: Transnational Connections and Local Social Realities* (London: Routledge, 2015).

Ng, Kim Chew, 'Sinophone/Chinese: "The South where Language is lost" and Reinvented, translated by Brian Bernards', in Shu-mei Shih, Chien-hsin Tsai and Brian Bernards (eds.) *Sinophone Studies: A Critical Reader* (New York: Columbia University Press, 2013), 74–92.

Sai Siew-Min, 'Mandarin Lessons: Modernity, Colonialism and Chinese Cultural Nationalism in the Dutch East Indies, c.1900s', *Inter-Asia Cultural Studies*, 17/3 (2016): 375–94.

Shih Shu-mei, 2004. 'Global Literature and the Technologies of Recognition', *Publications of the Modern Language Association of America*, 119/1: 16–30.

Shih Shu-mei, *Visuality and Identity: Sinophone Articulations across the Pacific* (Berkeley: University of California Press, 2007).

Shih Shu-mei, 'Introduction: What Is Sinophone Studies?', in Shih Shu-mei, Chien-Hsin Tsai and Brian Bernards (eds.), *Sinophone Studies: A Critical Reader* (New York: Columbia University Press, 2013): 1–16.

Stenberg, Josh, 'The Lost Keychain? Contemporary Chinese-Language Writing in Indonesia', *SOJOURN: Journal of Social Issues in Southeast Asia*, 32/3 (2017): 634–68.

Tam, Gina Anne, *Dialect and Nationalism in China, 1860–1960* (Cambridge: Cambridge University Press, 2020).

Tan, E.K., *Rethinking Chineseness: Translational Sinophone Identities in the Nanyang Literary World* (Amherst: Cambria Press, 2013).

Tsu Jing and David Der-wei Wang (eds.), *Global Chinese Literature: Critical Essays* (Leiden: Brill, 2010).

Tsu Jing, *Sound and Script in Chinese Diaspora* (Cambridge, Massachusetts: Harvard University Press, 2010).

Wang Xiaomei, *Mandarin Spread in Malaysia* (Kuala Lumpur: Oxford University Press, 2012).

Wu, Chia-rong, *Supernatural Sinophone Taiwan and Beyond* (Amherst: Cambria Press, 2016).

From *Ang moh* 紅毛 to *Phi jun* 批准:
The Role of Southern Min in Early Contacts between Chinese and European Languages

Picus Sizhi Ding

1 Introduction[1]

Literacy is usually considered to symbolize a high extent of civilization. Through texts, an enormous amount of knowledge can be maintained and transmitted properly. Texts also leave behind a record of language use over time. Nonetheless, a writing system is not a necessary phase of language development. In fact, the vast majority of human languages lack a writing tradition. This also holds true among Sinitic languages, where a northern prestige variety written in an ideographic system had represented for centuries the written standard, divorced from spoken Sinitic languages until the advent of a Mandarin-based national language in the early 1900s (see Chen 1999 and Hoogervorst, this volume). As such, it is often difficult to investigate issues related to sociohistorical linguistics in other Chinese varieties such as Southern Min, whose informal written form is found only in special genres, such as traditional opera scripts and other performance-based vocal works (van der Loon 1991; Chia, this volume). These texts were essentially created to cater the need of performers who would entertain an audience in front of them.

Lyrics – especially those for pop songs – represent an important additional source for written Cantonese and written Southern Min. These contemporary texts are extremely productive by virtue of the thriving entertainment industry in Hong Kong and Taiwan. In what follows I will first address lyrics of Cantonese pop songs, and later that of Southern Min. In the study of these and other non-Mandarin Sinitic texts, I propose to distinguish between literary-style texts and colloquial-style texts. The former are approximated to

---

1 In this chapter, all Chinese characters are presented in the traditional form with tone marks omitted in their romanization. Unless noted otherwise, the romanization of Chinese characters is based on Standard Chinese. The *Hanyu pinyin* 漢語拼音 and *Jyutping* 粵拼 are adopted for the romanization of Standard Chinese/Mandarin and Cantonese respectively. An informal scheme is used to transcribe Southern Min and Hakka.

the standard conventions of written Chinese, whereas the latter duly represent the written form of a Sinitic vernacular. Using excerpts of two Cantonese pop songs written by contemporary lyricists from Hong Kong, these two styles of texts are exemplified in (1) and (2) respectively. Sung in Cantonese, the songs belong to the same genre of vocal entertainment found in historical Sinitic texts in the south.

(1)   Literary-style Cantonese
      (a)   小小的   宇宙   天真的   宇宙
            siu siu dik   jyu zau   tin zan dik   jyu zau
            small   universe   innocent   universe
            'In the small and innocent universe'

      (b)   真的   我   真的   你   唔係   小木偶
            zan dik   ngo   zan dik   nei   m hai   siu muk ngau
            real   me   real   you   not.be   small.puppet
            'you and I are real, not small puppets'
            (Extracted from *Siu Si Hau* 小時候 'When We Are Small')

(2)   Colloquial-style Cantonese
      (a)   我哋   呢班   打工仔
            ngo dei   ni baan   daa gung zai
            we   this.group   blue-collar
            'We, being a group of workers,'

      (b)   通街   走糴   直頭係   壞腸胃
            tung gaai   zau dek   zik tau hai   waai coeng wai
            through.street   run.about   certainly   bad.belly
            'certainly (get) stomach sick by rushing from street to street'

      (c)   搵   嗰些少   到   月底   點夠   使
            wan   go se siu   dou   jyut dai   dim gau   sai
            find   that.little   till   month.end   how.enough   use
            'earning that little, enough for spending till end of the month?'
            (From *Bun Gan Baat Loeng* 半斤八両 'Half a Catty vs. Eight Taels')

It must be noted that style-switching (or borrowing between the standard and a vernacular variety) is a common phenomenon in non-Mandarin Sinitic texts. For instance, the Cantonese negative copula *m hai* 唔係 is found in (1b), an example of a literary style. This particular instance of style-switching is

necessary in order to accommodate a constraint on Cantonese lexical tones with musical tunes. In this regard, the lyric of *Bun Gan Baat Loeng* 半斤八両 'Half a Catty vs. Eight Taels', which was released in 1976, is remarkable in its utter avoidance of Standard Chinese expressions. In consequence, it perfectly reflects spoken Cantonese when the lyric is sung or read out. Such a degree of matching between writing and speaking is impossible with a literary-style text.

Nowadays in Hong Kong, lyricists of Cantonese pop songs often choose to write in the literary style, as doing so represents a higher register with a tincture of learned fashion. On the other hand, the Cantonese version of Wikipedia (zh-yue.wikipedia.org) adopts the colloquial style. Similarly, the Southern Min version of Wikipedia (zh-min-nan.wikipedia.org) features the colloquial style, but it is written in the Latin alphabet. By contrast, lyrics of contemporary Taiwanese (a variety of Southern Min) pop songs are written with Chinese characters. As the convention for using characters to write Southern Min is underdeveloped, the writing style found in most Taiwanese lyrics holds somewhat of a middle ground between the literary and colloquial style.[2]

Examples (3) and (4) below present excerpts of lyrics displaying the two styles of Southern Min texts. The lyrics written in the literary style in (3) are taken from a duet in which a woman sings in Mandarin and a man in Taiwanese. This unusual combination of languages in the song has probably prompted the choice of a style closer to Mandarin (i.e. the literary style). In colloquial-style writing, the sound principle – that is, the approximation of Southern Min pronunciation – represents a favoured strategy for the selection of Chinese characters for a word, e.g. *phwey* 批 instead of *sin* 信 (which would be taken under a meaning-based approach) for 'letter', as found in the title of the song in (4). This principle for character selection is observed much more consistently in written Cantonese than in written Southern Min. As shown in (4), it is not uncommon to find varying preference over the choice of characters for the same word by different people in colloquial-style Taiwanese pop songs, e.g. 乎/予 for *hoh* 'to give' and 多/濟 for *jwey* 'much, many'.

(3)   Literary-style Taiwanese

    (a)   為   妳   傷過   的   心

           wey   li   siong kwey   ey   sim

           for   you   hurt.EXP   of   heart

           'The heart that has been hurt because of you'

---

2   For a general review of the varied orthography of Southern Min, see Ding (2016: 70).

(b)

孤單	望	月	又	一	年
koh tan	mang	kwey	iu	jit	nin
lone	look	moon	again	one	year

'look at the moon alone for another year'

(c)

不甘	離開	有	妳	的	記誌
m kam	li khwi	u	li	ey	ki ti
reluctant	leave	have	you	of	memory

'unwilling to go off the memory of you'
(From *Kim Sin Ji Wey Li* 今生只為你 'This Life for You Only')

(4) Colloquial-style Taiwanese

(a)

感謝	你	乎/予	我	愛過
kam sia	li	hoh	kwa	ai kwey
thank	you	give	me	love.EXP

'Thank you for having let me love you'

(b)

乎/予	我	心痛	也	這	多/濟
hoh	kwa	sim thia	ya	jia	jwey
give	me	heartache	also	this	much

'giving me also so much heartache'

(c)

恨	我	自己/甲己	袂	覺醒
hun	kwa	ka yi	wey	kak chin
hate	me	oneself	cannot	awaken

'I hate myself for being unable to wake up'

(d)

無	資格	會凍/會當	講	後悔
mo	chu keh	ey thang	kong	ho hwey
not.have	capability	can	say	regret

'have no right to express regret'
(From *Jit Tiong Phwey* 一張批 'A Letter')

Taking Cantonese as an example, Table 1.1 summarizes the major differences of these two styles of texts. The literary style, with its vocabulary and grammar largely shared with Standard Chinese,[3] is so close to writing in Standard Chinese that it can be considered a variety of written Chinese, readable by

---

3   This refers to contemporary Mandarin in modern times and Classical Chinese (文言 *wenyan*) in the past.

TABLE 1.1   Literary-style Cantonese versus colloquial-style Cantonese in writing

	Literary-style Cantonese	**Colloquial-style Cantonese**
Vocabulary	Largely from Standard Chinese	As a norm, Cantonese words
Grammar	Based on Standard Chinese	Based on Cantonese
Intelligibility	Barring Cantonese expressions, intelligible to all Sinitic speakers	Barely intelligible to other speakers of Sinitic

everyone literate in Chinese. In contrast, the colloquial-style text is difficult to comprehend for those who do not speak the particular vernacular. Because of this distinction, the aforementioned two styles of non-Mandarin Sinitic texts can be identified.

This distinction is relevant to the rest of this chapter. Historical texts written in Southern Min appear to be of the literary style, or a mix of both styles. In *Nai Kia Ki* 荔鏡記 'Tale of the Lychee Mirror', which is dated to the sixteenth century and regarded as the first written literature in Southern Min, the general distribution pattern of writing styles is as follows: literary style for narratives (which read like Classical Chinese), and colloquial style for dialogues (see also Chia, this volume). As a result, Southern Min texts of early-modern times that faithfully represent its vocabulary and grammar are few and far between. Identifying a Southern Min word or expression written in Chinese characters (sometimes even when it is romanized, as in Pidgin English) is usually not a straightforward matter. Nevertheless, doing so enables us to unearth a number of elusive trajectories of language contact that form the backbone of this chapter.

In spite of the challenges posed by the meagre quantity of texts available for Southern Min, complicated by the issue of writing style, this chapter attempts to discuss some lexical contributions from Southern Min to Sinitic as a whole, underscoring its role in early Sino-European contact and international trade in Southeast Asia and southern China. These include the spread of the terms *ang moh* 紅毛 'European; white person', *toh lien* 榴槤 'durian', *ang moh tan* 紅毛丹 'rambutan', and *han ji* 番薯 'sweet potato' from Southern Min to other Sinitic languages, and a theory postulating *phi jun* 批准 from Southern Min as the source of *pidgin* in English. This chapter pays particular attention to *ang moh* 紅毛, once a popular loanword found in many Sinitic languages and attested in old Chinese texts, and *phi jun* 批准. It offers a solution to the obscure origins of *pidgin* against the backdrop of significant contact between Southern Min and English since the early history of Sino-British trade.

This chapter is organized as follows: a brief ethnographic account of the Hokkien people and their languages is presented in the next section. Then the term *ang moh* 紅毛 is studied in detail, to be ensued by discussions of *han ji* 番薯 'sweet potato', *toh lien* 榴槤 'durian', and *ang moh tan* 紅毛丹 'rambutan' as additional instances of lexical diffusion from Southern Min to other Sinitic languages. A large section is then devoted to unveiling the origin of *pidgin*. Firstly, the essential role of Hokkien merchants held in centuries-long Chinese foreign trade is addressed, and connections between Southern Min and Chinese Pidgin English are studied. Then, a morphological process known as clipping is scrutinized for its different applications in Mandarin, Cantonese and Southern Min. A hypothesis is advanced to account for the semantic change and borrowing of *phi jun* 批准 into English. Finally, this chapter concludes with remarks on these lexical contributions from Southern Min as the outcome of its intensive contact with foreign languages, including those of Europe as well as Southeast Asia.

2 Hokkien: An Ethnographic Outline

The term *Hokkien* 福建 comes from Southern Min and refers to the province of Fujian. In English usage, this toponym is also employed as the name for the major language spoken in southern Fujian: Southern Min. In this sense, *Hokkien* can be regarded to be a synonym to Southern Min, yet it must be emphasized that Fujian is one of the most linguistically diverse provinces in the coastal region of China, where a variety of mutually unintelligible Sinitic languages are spoken (Norman 1988). These include the Min group, as shown in Fig. 1.1 (based on Wurm et al. 1988). In addition, Hakka and Mandarin are also spoken in Fujian province. Among these Sinitic languages, Southern Min has the largest population and the widest distribution in southeast China. Smaller Min varieties such as Pu-Xian 莆仙, Leizhou 雷州, and Hainan 海南 have their roots in Old Southern Min, since their ancestors settled in or launched sail from the then Southern Min-speaking area (Li & Yao 2008). In modern times, however, their mutual intelligibility with Southern Min has been lost, and therefore, they have an independent status within the Min group.

Situated on the west side of the Taiwan Strait in south-eastern mainland China, Fujian is characterized by extensive mountainous and rugged terrain covering 95% of its land; four small plains lie on the east and southeast coast (Tang 1995: 27). The hills, however, provide excellent plantations for tea cultivation, and famous varieties of oolong tea are found in the Wuyi Mountains 武夷山 in the north and Anxi 安溪 in the south, rendering Fujian one of the

most important tea-producing centres in China. Indeed the word for *tea* in English, Dutch, Malay, and many other languages can be traced to *tey* 茶 in Southern Min (even if the word is not directly borrowed from Southern Min). Fujian also has China's largest proportion of forestland, consisting of over 60% of land in the province. A high concentration of settlements along its shores has led to an extremely dense population with a resource scarcity of arable land. Consequently, Southern Min people have adopted a maritime orientation for living: "taking the sea as their farming field" (Li & Yao 2008: 140).

Mazu 媽祖, also known as Lin Mo Niang 林默娘 during her life, was a native of Fujian from an isle in Putian 莆田 who became exalted as the goddess of seafarers and fishermen. The worship of *Mazu* is shared among the linguistically diverse Min people as well as non-Min speaking Chinese living on the shores. *Mazu* temples, also known as "heavenly queen temples" 天后廟,

are frequently found in Min communities, not only in mainland China and Taiwan, but also in a large number of cities in Southeast Asia (see Xu 2007). The excellent maritime navigation skills of the Southern Min people have facilitated the development of harbours and seaports in Quanzhou 泉州 and Zhangzhou 漳州 in southern Fujian. During the Song and the Yuan Dynasties (ca. 1200s–1350s), Quanzhou served as the starting point of the Maritime Silk Route, which connected southern Fujian to Southeast Asia and beyond via the South China Sea, the Straits of Malacca, and the Java Sea. Since the late-fourteenth century, a closed-door policy prohibiting maritime trade came into force repeatedly during the Ming and the Qing Dynasties. In 1567, when the Ming court decided to resume limited entrepôt trade with foreign countries, Zhangzhou was selected as the designated seaport. In the wake of the Sino-British Treaty of Nanking of 1842, Amoy 廈門 swiftly arose to its leading role of entrepôt in the Southern Min region (Wang 2011).[4]

Centuries of trade via the Maritime Silk Route have resulted in the establishment of Chinese settlements in various parts of Southeast Asia (Tamura et al. 1997: 70; Miksic 2013: 20; Lee 2013). Natives of southern Fujian who migrated to Southeast Asia between the 1500s and 1940s typically possessed the following characteristics: (a) being a monolingual speaker of Hokkien, (b) believing in *Mazu*, and (c) engaging in business, such as wholesale or retail trade.[5] The majority of Southern Min communities in Southeast Asia have maintained their mother tongue until unfavourable language policies were imposed by governments in the past four decades or so. Under these circumstances, language shifts from Southern Min to a national language precipitated the loss of Hokkien to Burmese in Myanmar, to Indonesian in Indonesia, and to English in Singapore. In the latter country, Mandarin is designated as the "heritage language" for ethnic Chinese regardless of their origin (Ding 2016).

3 *Ang moh* 紅毛: Contact between Hokkien and European Languages

At the turn of the sixteenth century, during the Age of Discovery in European history, waves of Europeans reached the Far East and established trading ports in Southeast Asia, first led by the Portuguese and Spaniards and then followed by the Dutch, British and French. Fig. 1.2 illustrates the new sea route navigated by Portuguese explorers, who first visited Malacca in 1509 by way of

---

4 In addition to Amoy, Canton, Fuzhou, Ningbo and Shanghai were all open to trading with Europeans in this period.

5 See Ng (2015) for a further study of the commercial orientation of Hokkien speakers.
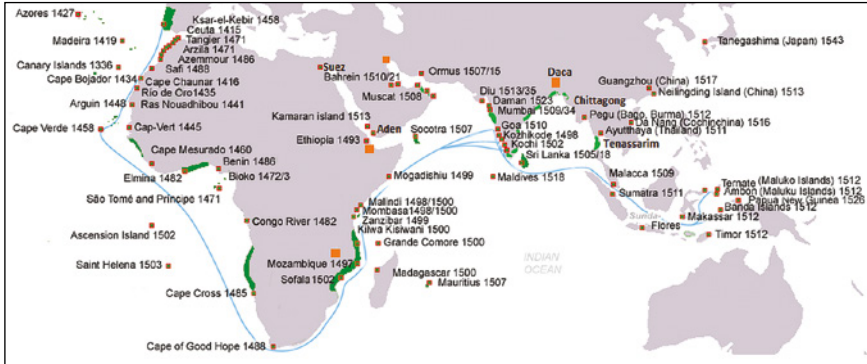
FIGURE 1.2   The sea route from Portugal to the Far East

the Cape of Good Hope and southern India, and eventually penetrated South China upon reaching the waters of the Pearl River in 1513.

It is clear from European accounts that a metropolis had developed in Malacca – where the Chinese, Japanese, and Arabs met for trade over a lengthy period (Lach 1994: 501) – prior to the advent of Portuguese fleets in 1509. A small Hokkien community emerged in Malacca no later than the early fifteenth century (Lee 2013: 406–7). According to Chinese historical accounts (Ma 1451), Chinese communities consisting of natives of southern Hokkien as well as natives of Guangdong were found in various islands of present-day Indonesia, including Sumatra and Java, during Zheng He's 鄭和 visits to Southeast Asia in the early 1400s (see also Lee 2013: 142–43).[6] With decades of unceasing waves of immigration from southern Fujian, the Chinese traders grew into a large community, while their language, Southern Min, became the dominant language of the Southeast Asian Chinese.[7]

Similarly, a considerable Chinese population, the majority of whom were Hokkiens, had resided in Indonesia well before a fleet of four Dutch ships first reached the island of Java in 1596.[8] The Dutch East India Company (VOC) was

6   Known as the Ming treasure voyages, seven maritime expeditions were undertaken by Ming China's treasure fleet between 1405 and 1433, reaching a number of countries in Southeast Asia, Middle East and East Africa (for details, see Levathes 1996).

7   See Aye (this volume) for an argument that Baba/Bazaar Malay evolved from a Hokkien-Malay pidgin. For a history of ethnic Chinese of Southeast Asia, see Lee (2013).

8   See "The Dutch East India Company's shipping between the Netherlands and Asia 1595–1795" (retrieved on March 25, 2020): http://resources.huygens.knaw.nl/das/voyages.

established in 1602 to safeguard Dutch commercial interests in the Far East
(Walker 2012: 315). During the first half of the seventeenth century, the Dutch,
under the auspices of the Dutch East India Company, emerged as a leading
European power by establishing a number of trading ports in Southeast Asia
and Taiwan (Andrade 2007; De Witt 2009). Permanent Dutch trading posts in
Indonesia were founded in northwest Java: the first one in 1603 and the sec-
ond one, also the most important one, in Jayakarta (nowadays Jakarta) in 1611.
The 1620s witnessed a series of military conflicts, the "Sino-Dutch battles",
when the Dutch intention to open a trading port on the shores of Fujian was
resisted by the Chinese court (Twitchett & Mote 1998). The Dutch, despite
being defeated in the war, were granted permission to exploit the island
Taiwan, then regarded to be an offshore frontier of China. After the island's
Austronesian-speaking aboriginals had been pacified or driven out of the
plains in the 1640s, "Dutch Formosa" (1624–62) launched a large campaign to
entice natives of southern Fujian to immigrate to the island. In the meantime,
Dutch Malacca replaced Portuguese Malacca on the Malay Peninsula by 1641.

Sino-Dutch encounters continued in later centuries. In the mid-nineteenth
century, the colonial government requested and supported the training of the
first cohort of Dutch Sinologists to be employed as consultants stationed at
various Indonesian ports. An emphasis was placed on their ability to speak
Southern Min, so that they could assist in the management of local affairs in
this Dutch colony (Kuiper 2017).

These encounters raise the question how the Dutch were designated by
Hokkien speakers. In present-day Southern Min, especially its Southeast Asian
varieties (see Churchman, this volume), the epithet *ang moh* 紅毛 (lit. 'red
hair'), means 'European' or more generally 'white person; Caucasian'. In terms
of genetic traits, individuals with red or auburn hair are much more common
in the Dutch population than among the Portuguese.[9] Apparently the Hokkien
living in Indonesia used this epithet to refer to this new group of Europeans,
who were physically distinguishable from the earlier Portuguese explorers,
by the late 1590s.[10] This semantic expansion reflected a successive shift of

9   A map showing distribution of light hair in Europe can be found at the following link
       (retrieved on August 2, 2018): http://unsafeharbour.wordpress.com/2012/02/19/distribu
       tion-of-light-hair-and-eyes-in-europe/.

10   Red hair was often portrayed for Dutchmen in Chinese and Japanese paintings in the
       1700s and later. For examples (retrieved on March 28, 2020), Zhang Rulin's work dated
       to 1738, http://www.rijksmuseum.nl/en/collection/AK-MAK-1410, and Kawahara Keiga's
       drawing of 1811, http://navstory.co/dejima-dutch-ship (available also in commons.wiki
       media.org).

European powers in Southeast Asia. Subsequent to the Dutch, other European imperial powers also wielded their influence in this region and came in increasing contact with the Hokkien. These new Europeans are all known as *ang moh* 紅毛 in Southern Min, regardless of their nationality. It is therefore possible to interpret the essential meaning of this epithet as 'blond-haired white people (from Europe)', irrespective of nationality. For this reason, the English missionary Walter Henry Medhurst, having studied Hokkien in British Malaya, glossed the word *ang moh* 紅毛 as 'English people' in his *Dictionary of the Hok-Këèn Dialect of the Chinese Language* (Medhurst 1832: 481). This semantic shift occurred in all Hokkien varieties spoken in Southeast Asia.

3.1 *The Spread of* Ang moh 紅毛 *to Other Sinitic Languages*
A Ming-dynasty travel writing titled *Yue Jian Bian* 粵劍編 (lit. 'Guangdong sword book') represents one of the earliest available texts where the word *ang moh* 紅毛 is attested in a compound. This book was authored by a mandarin named Wang Linheng 王臨亨 (1556–1603), who visited southern Guangdong on an official trip from Zhejiang in 1601. The following excerpt from the book suggests one of the first Sino-Dutch contacts within the territory of China:[11]

> 辛丑九月間，有二夷舟至香山澳，通事者亦不知何國人，人呼之為**紅毛鬼**。

> Around the ninth lunar month in 1601, two foreign boats arrived at Macao. Even the interpreter did not know of their nationality, and people called them *red-haired ghosts*.

This brief mentioning of the first Chinese encounter with the Dutch in Macao (which by then was on lease to the Portuguese) corroborates the origins of the term *ang moh* 紅毛 in Southern Min. As noted above, Indonesia was gradually transformed into a Dutch colony in the wake of their first landing in 1596. Well before that, the Hokkien had settled there in significant numbers. The above passage clearly indicates that the word *ang moh* 紅毛 predated this particular episode of Sino-Dutch contact in Macao, even though the interpreter specialized in dealing with Europeans was not aware of the home country of these new visitors. In other words, some local Chinese from the south must have employed this epithet to refer to the Dutch by 1601, most likely through

---

11   For details, see http://ctext.org/wiki.pl?if=gb&res=187407&searchu=紅毛.

contact between Indonesian Hokkien, Cantonese, Hakka, or other Sinitic varieties used among Chinese communities in Southeast Asia.[12]

The word 紅毛 (Mandarin: *hong mao*) appeared alongside its formal equivalent 紅夷 (*hong yi*; lit. 'red barbarian') in Yu Yonghe's 郁永河 undated *A Diary of Sulfur Mining* 採硫日記, which narrates a business trip taken in 1697 from Fuzhou 福州, the capital of Fujian, to Taiwan. In addition, 和蘭 (*he lan*), a transliteration of *Holland*, was adopted in formal historical writing compiled in a later time such as the *Ming Shi* 明史 'Chronicle of the Ming Dynasty', e.g. 和蘭又名紅毛番 "Holland, alias, land of red-haired barbarians" (scroll 325).[13] This formal term 和蘭 (*he lan*) was used less commonly than 紅毛 (*hong mao*) in written Chinese. Other written records in which the latter occurred include official documents of the mid-1700s and titles of self-taught booklets for learning Pidgin English in the 1800s (see Fig. 1.4 in section 5).[14] These texts evince that the rate of diffusion from Southern Min to other Sinitic languages was rapid, suggesting vigorous activities and contacts with the Chinese in Southeast Asia and South China since the Dutch arrival in the Far East in 1596.

Southern Min has derived many compounds from *ang moh* 紅毛, e.g. *ang moh tey* 紅毛茶 'red tea' (lit. 'red-hair tea'), *ang moh wey* 紅毛話 'English' (lit. 'red-hair speech'), *ang moh hey* 紅毛灰 'cement' (lit. 'red-hair dust') or *ang moh toh* 紅毛塗 'cement' (lit. 'red-hair mud'; mainly in Taiwanese), and *ang moh jun* 紅毛船 'European ship' (lit. 'red-hair ship'). In present-day Cantonese, while the term *hong mou* 紅毛 does not occur by itself, it is found in the compound *hong mou nai* 紅毛泥 'cement' (lit. 'red-hair soil'), which is still commonly used alongside the Standard Chinese term *shui ni* 水泥 (lit. 'water soil'). In Hakka, especially the varieties spoken in Taiwan, cement is known as *fung mo foi* 紅毛灰 (lit. 'red-hair dust') or *fung mo nai* 紅毛泥 (lit. 'red-hair soil').

Nevertheless, 紅毛 *hong mao* and its derivations, in both written and spoken forms, have entirely retreated from standard Mandarin in mainland China since the completion of its lexical renovation, a process in which loanwords or vernacular forms were replaced by new preferable terminology. For instance, the terms *da ge da* 大哥大 'cell phone' (lit. 'big brother big') and *fei lao* 肺癆 'pulmonary tuberculosis' (lit. 'lung tuberculosis') are nowadays archaic in

---

12　Reviewers of this chapter noted that the characters 紅毛 are found written on a seventeenth-century map, which depicted the Maluku Islands in eastern Indonesia (http://www.straitstimes.com/singapore/term-ang-moh-in-use-as-early-as-1600s-in -ming-dynasty-map).

13　http://zh.wikisource.org/wiki/明史_(四庫全書本)/卷325#明史卷三百二十五 (retrieved on March 21, 2020).

14　http://zh.wikisource.org/wiki/皇朝文獻通考_(四庫全書本)/卷033 (retrieved on August 4, 2018).

Standard Chinese in mainland China (but still used in other Sinitic varieties). The former has been replaced by *shou ji* 手機 (lit. 'hand device') and the latter by *fei jiehe* 肺結核 (lit. 'lung nodule').

3.2 *The Borrowing of* Ang moh 紅毛 *into Japanese*

Through the use of *kanji* in the Japanese writing system, the term 紅毛 (read in Japanese as *kōmō*) was borrowed into Japanese, referring primarily to the Dutch and being restricted to the Edo period (1603–1868). This is a likely instance of borrowing from Southern Min, as the Japanese word for *red* is written with a different *kanji* 赤, as found in Old Chinese texts, cf. 赤毛 (*aka ge*) 'redhead'. The Japanese scholar Arai (1715) noted that the Chinese called the Dutch people 紅毛 (*hong mao*, *ang moh*), but no comment was made regarding its etymology in Japanese. In contrast to the lack of differentiation between the Portuguese and the Dutch in Chinese languages (at least in the vernacular languages),[15] these two groups of Europeans were distinguished in Japanese since the early days of contacts: the Portuguese were known as *namban* (南蠻, or 南蛮 in present-day simplified *kanji*; lit. 'south barbarian'), whereas the Dutch were designated as *kōmō* 紅毛 (lit. 'red hair'). To some extent, these words can be expanded, e.g. *namban* 南蠻 may also refer to 'Spaniard', especially in the compound *namban bōeki* 南蠻貿易 'trade with the Portuguese and/or Spaniards', and *kōmō* 紅毛 may cover 'Occidentals' in a vague sense.

The first Luso-Japanese contact occurred in September 1543, when a Portuguese boat originally headed for Ningbo 寧波 (situated on China's eastern coast) was carried by a powerful storm to Tanegashima, a small island offshore southern Kyushu (Kshetry 2008: 39). According to a Japanese account recorded in *Teppōki* 鐵炮記 (lit. 'iron cannon log'),[16] a Chinese crew member from the boat communicated with the Japanese by writing Chinese on the sand, introducing the Portuguese as *xi nan man zhong* 西南蠻種 (lit. 'southwest barbarian sort'), presumably because they sailed from their foothold in Malacca. Based on this, the Japanese appropriated the term *namban* 南蠻 to refer to the Portuguese. Nonetheless, the Chinese have never applied this term to any Europeans.

---

15   In formal historical Chinese texts such as the *Ming Shi* 明史 'Chronicle of the Ming Dynasty', the Portuguese were referred to with the term *fo lang ji* 佛郎機, based on *franjī* 'Franks' in Arabic. In less formal writing, they were called *xiang shan ao yi* 香山澳夷 (lit. 'barbarians of Macao'), or abbreviated to *ao yi* 澳夷. None of these terms were used in Cantonese or Southern Min, however.

16   This text appeared in 1601, almost six decades following the incident. For the original source, see https://ja.wikisource.org/wiki/鉄炮記 (retrieved on March 18, 2020).

It is also by accident that the first Dutch ship, *de Liefde* (Love), reached Bungo on the eastern shore of Kyushu in April 1600, having been blown off its original course to the Spice Islands (Maluku, Indonesia) by a violent storm (Goodman 2000: 9). Yet in this case, there was no Chinese person on the ship, which had departed from Holland. Instead, there was an Englishman, William Adams, serving as the pilot. Alongside the Dutch sailor Jan Joosten, the two became the first Europeans granted the title of *samurai* by a *shogun*, and they settled in Japan with their new Japanese names: Miura Anjin 三浦按針 (lit. 'the pilot of Miura') for William Adams and Yayōsu 耶楊子 for Jan Joosten. No historical account is available on the way the Japanese called the Dutch after their initial contact in Kyushu.

Japanese terminology for the Netherlands consists of a pair of synonyms: *oranda* 阿蘭陀 ~ 和蘭陀 versus *kōmō* 紅毛. The former represents a higher register, as it stems from a transliteration of the autonym *Holland*, while the latter was probably an epithet borrowed from Southern Min through written Chinese. The Ryukyu Kingdom (which became the Okinawa Prefecture after the annexation of Japan in 1879), being both a tributary state of China and a vassal of Japan, held a key role in maritime trade extending from Northeast Asia to Southeast Asia in the fifteenth and sixteenth centuries (Akamine 2017). Its function as a bridge between Japan and Ming China in this period was especially vital to the Japanese in terms of conducting indirect trade with China and receiving information on Chinese society, since the Sino-Japanese relationship had been severed in this period. Therefore, a probable borrowing channel of the term 紅毛 *kōmō* into Japanese would be in its written form through Ryukyuan traders, in an extended chain of borrowing: Southern Min > Written Chinese > Ryukyuan > Japanese.

3.3　*A Synopsis of the Spread of* **Ang moh** 紅毛

Following a general pattern of European contact with China by maritime routes, the Dutch first encountered the Chinese people outside China. More specifically, informal Sino-Dutch contacts began in Indonesia, where the Dutch met Hokkien-speaking Chinese in 1596 or shortly afterwards. This contact led to the creation of the epithetic expression *ang moh* 紅毛 in Hokkien. To conclude this section, the spread and development of *ang moh* 紅毛 is outlined in Fig. 1.3.

4　　Diffusion of Southern Min Words into Other Sinitic Languages

Following the promotion of Mandarin in Taiwan by the Chinese Nationalist Party (國民黨 Kuomintang), especially under the national language campaign
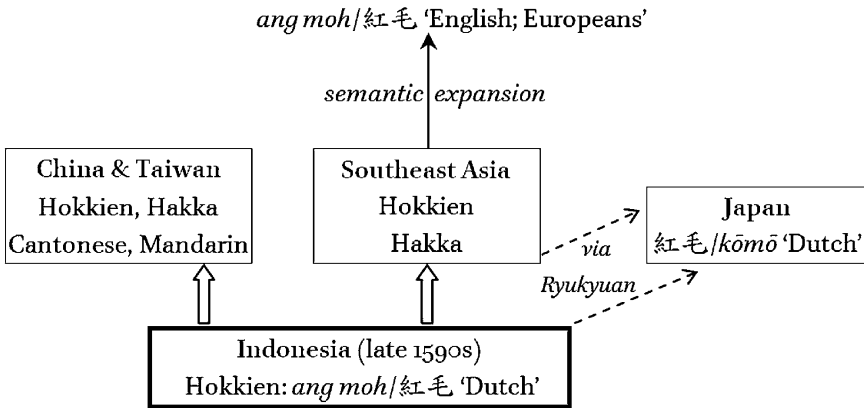
FIGURE 1.3   The spread and development of ang moh 紅毛
OWN WORK

in the mid-1940s (Ding 2016: 64–68), a number of loanwords have entered Taiwanese Mandarin from Southern Min owing to intensive language contact. Examples include Southern Min *mo jua* 'without difference; does not matter' rendered in Mandarin as *mei cha* 没差, *sam pat* '(derogatory) a spiteful or unpleasant woman' as *san ba* 三八, and *swi jia boh* 'pretty woman' as *shui cha mou* 水查某. Barring such examples limited to Taiwan, lexical diffusion from Southern Min to other Sinitic languages appears to be extremely rare in the Sinitic heartlands except for an emerging trend in Internet literature.

Dispersed across Southeast Asia from premodern times, Southern Min varieties have long been exposed to a number of tropical products. Naturally, some fruit and vegetable names thus entered Standard Chinese and other Sinitic languages via Hokkien, although such etymologies are often not realized by the speech community. Three instances of this are *han ji* 番薯 'sweet potato', *toh lien* 榴槤 'durian', and *ang moh tan* 紅毛丹 'rambutan'. As with *ang moh* 紅毛, lexical diffusion of these words occurred smoothly by means of Chinese characters. In fact, this has obscured their etymology due to the loss of sound associations in the process of character-based infusion.

The sweet potato (*Ipomoea batatas*) is currently one of the most widely grown crops in China. It is called *han ji* in Hokkien and *faan syu* in Cantonese, written identically as 番薯 (lit. 'foreign tuber') in both languages. In other parts of mainland China, the sweet potato is known by scores of terms on account of lexical renovation in different varieties of Sinitic, e.g. *hongshu* 紅薯 (lit. 'red tuber'), *ganshu* 甘薯 (lit. 'sweet tuber'), and *shanshu* 山薯 (lit. 'mountain tuber') (Ni & Xiang 2014). As such, it is not possible to identify the first designation of the sweet potato in the Sinitic family solely on linguistic

clues, but historical evidence may help us further. According to Ho (1955) and Simoons (1991: 102), in 1594, the governor of Fujian, Chin Hsüeh-tseng 金學曾, urged farmers to grow sweet potatoes in response to crop failures caused by natural disasters. This signified the first campaign leading to its large-scale cultivation in China after Chen Zhenlong 陳振龍, a native of Fujian, brought sweet potatoes from Luzon in the Philippines a few years earlier. The cultivation of sweet potatoes subsequently spread from coastal Fujian to eastern China and beyond (Ni & Xiang 2014; Zhang et al. 2009). In light of the strong ties between southern Fujian and Southeast Asia, Southern Min is the most likely Sinitic language from which the sweet potato received its initial name in China.

In recent decades, the durian (*Durio zibethinus*) has become rather popular in mainland China, being regularly sold at supermarkets in many cities. Misguided by the Chinese characters 榴槤 (Mandarin: *liu lian*, Hakka: *liu lien*, Cantonese: *lau lin*), few Chinese recognize its Malay etymology. Interestingly, Ma Huan 馬歡 (1451) mentioned the term 賭尔鳥 (*du er niao* in modern Mandarin, cf. *durian* in Malay) in reference to an exotic fruit of Sumatra with a compelling smell.[17] The choice of the third character was a rather creative attempt to represent a nasalized rhyme; in conjunction with the consonant indicated by the second character, the two would have denoted the sound *riã*, using the centuries-long Chinese phonological tradition of *qie yun* 切韻 (lit. 'cut rhyme'). Despite its approximation of the word's Malay pronunciation, this transliteration did not gain traction in Chinese, since it was infeasible to export this fruit as a commodity to fifteenth-century China. In Burmese Hokkien, the durian is still called *toh lien* 榴槤, which is much closer to the Malay form *durian* than its pronunciation in Cantonese, Hakka, or Mandarin. This would indicate that the Malay loanword has spread from Southeast Asian Hokkien to other Sinitic languages through Chinese characters, which do not necessarily reflect its original pronunciation in the donor language. It is also worth pointing out that in the literary reading of Southern Min for 榴槤 *liu lien*, the dental initial of the first syllable becomes a lateral; such a sound correspondence is regularly observed in Southern Min.[18]

The rambutan (*Nephelium lappaceum*) is another tropical fruit from the Malaysian-Indonesian region. Known in Chinese as 紅毛丹 (Mandarin: *hong*

---

17   In later editions of the texts, the original characters 賭尔鳥 (*du er niao*) were altered to 賭爾焉 (*du er yan*).

18   For example *twi* 鐳 'money' borrowed from Malay *duit* (which, in turn, is derived from Dutch *duit* 'a copper coin'). Examples of this sound correspondence in Philippine Hokkien can be found in Chia (this volume).

mao dan), Hokkien: *ang moh tan*, Cantonese: *hong mou daan*, Hakka: *fung mo dan*), the Chinese naming of this fruit appears to have been based on its appearance. In fact, the Malay name of this fruit, *rambutan*, likewise contains the word *rambut* 'hair'. Upon scrutiny, however, a puzzle arises from the choice of the third character 丹 'crimson', with the literal meaning of the Chinese term being 'red-hair crimson'.[19] If a descriptive approach had been taken in the naming of this fruit, it would have been 紅毛果 (*hong mao guo*; lit. 'red-hair fruit') or 紅毛荔枝 (*hong mao li zhi*; lit. 'red-hair lychee'). The apparent sound correspondence between the third syllable of Malay *rambutan* and Chinese 丹 – transcribed as *tan* or *da(a)n*, depending on the romanization scheme and variety of Sinitic – bespeaks a transliteration of the original Malay name in Chinese. Among the pronunciations of 紅毛丹 in the four major Sinitic languages – Hokkien: *ang moh tan*, Mandarin: *hong mao dan*, Cantonese: *hong mou daan*, and Hakka *fung mo dan* – Hokkien stands out as the best candidate for the role of recipient. In the case of Southern Min, the only phonological modification involved is the deletion of the rhotic initial in the Malay word, after which a near-perfect name emerged for this exotic fruit characterized by red hair on the skin.

5   *Phi jun* 批准: A Hokkien Hypothesis for the Origin of *Pidgin*

In the field of Pidgin and Creole Studies, the etymology of the term *pidgin* has remained unresolved since its inception. A common view, as noted in Oxford English Dictionary, holds that *pidgin* – or *pigeon*, as in "pigeon English", which was in use between 1859 and 1876 – is derived from a Chinese pronunciation of the English word "business". This is attributed to the fact that Chinese Pidgin English emerged out of communicative needs for Sino-European trade, hence the presumed connection between *pidgin* and *business*. However, there is no linguistic evidence to support this claim. Li et al. (2005) attempt to explain the etymology of *pidgin* through a Cantonese truncation of 'business' to 'busin', yielding *bit zin* 必剪 or *bei zin* 卑剪. Yet compared to the Cantonese borrowing of 'biscuit' as *bei si git* 卑士結 or 'bus' as *ba si* 巴士, the hypothetical rendering

---

19   Coincidentally, the previously mentioned term *da ge da* 大哥大 'cell phone' (lit. 'big brother big') seems to allow a redundant use of descriptive morpheme in the word. The actual morphological structure of this word consists of two parts: *dage* 'big brother' and *da* 'big'. Under the same analysis 紅毛丹 (*ang moh tan*) would be composed of *ang moh* 'Dutch; European' plus *tan* 'crimson', meaning 'something red from the Dutch or from Europe'.

of *bit zin* 必剪 or *bei zin* 卑剪 for 'business' is rather aberrant.[20] In fact, this etymology for *pidgin* was rejected as early as the 1870s, alongside other hypotheses such as Portuguese *ocupação* 'occupation' as its source (Leland 1876: 131; Holm 2000: 9; and references therein).

Belonging to the core vocabulary of Chinese Pidgin English, *pidgin* is a versatile word that covers many meanings and functions. Leland (1876: 131) listed 'business; affair; occupation' as its basic meanings, providing examples of its derivations such as *joss-pidgin* 'religion' and *chow-chow-pidgin* 'eating or cookery'. Since semantic extension takes time to accomplish, the word should exist in the mixed language decades before its first appearance in the mid-nineteenth century in the expression written as *pigeon English*. Hence, by the time *pidgin* induced substantial interest and attention from scholars, its etymology had been lost.

Throughout the history of expansion of the British Empire in the Far East, contact between English and Sinitic languages took place not only in mainland China, but also in Southeast Asia. Thus, the early formation of Chinese pidgin varieties of English should not be confined to the interaction between the English and Cantonese in the Pearl River Delta region, as has been the case in previous studies (cf. Li et al. 2005; Ansaldo et al. 2010). This assumption has restricted the scope of investigation: linguistically speaking, its focus on Cantonese has led to negligence of other varieties of Sinitic; geographically speaking, it fails to recognize Southeast Asia as a significant region for Anglo-Chinese contact.

To trace the etymological origin of *pidgin*, an understanding of the sociohistorical settings of British trade in the Far East is necessary. Despite being a relatively late arrival, the British Empire, following the signature of the Anglo-Dutch Treaty of 1824, successfully consolidated its colonies in Southeast Asia. Since then, a vigorous recruitment of Chinese from Fujian and Guangdong contributed to the rapid growth and reinforcement of Chinese communities in British Malaya, including Singapore (Song 1967; Cheng 1985; Lee 2013; Ng 2015). Consequently, Sinitic languages – Southern Min in particular – represented important parts of the linguistic landscape of colonial Southeast Asia. As such, the Malay Peninsula could be regarded the linguistic backyard of the Sinophone even prior to British rule.

---

20   Uchida (2009) provides many examples of Chinese Pidgin English, using Chinese characters to mark pronunciation for reading in Cantonese.

5.1   *Hokkiens as Major Chinese Traders in Canton and Southeast Asia*

Canton (nowadays Guangzhou 廣州) has been a major centre for Sino-European trade ever since it joined Zhangzhou in 1685 as one of three additional entrepôts on China's coast. The trade volume between Britain and China was accelerated by an insatiable demand of tea on the British market since the 1750s (Li 2010: 22). As a result, Britain swiftly assumed a top position in the Sino-European trade by the late eighteenth century. For instance, the total value of the Sino-British trade in 1792 stood more than 24 times that of the Sino-French trade. The expansion of Sino-British trade largely coincided with the Chinese adoption of the so-called "Canton Trade System" (1757–1842), which designated Canton as the sole port for receiving foreign traders travelling to China.[21] A number of restrictions and regulations were imposed on European merchants sojourning to Canton. For instance, foreigners were prohibited from learning Chinese or teaching foreign languages to the local population. Furthermore, no direct communication with the government was permitted. Instead, foreigners had to deal with their Chinese counterparts, *hong* 行 merchants (Downs & Grant Jr. 2014: 73–74).

As early as the turn of the eighteenth century, a number of merchants from Amoy decided to try their fortune in foreign trade in Canton. The following entrepreneurs, of considerable fame in Canton between the 1700s and 1730s, all hailed from southern Fujian: Limia, Anqua, Kimco, Shabang, Suqua, and Cowlo; each year they travelled regularly between Canton and Fujian (Ng 2015: 174; Van Dyke 2011: 80; Wang 2011, chapter 9). As a matter of fact, Hokkien businessmen dominated the international trade in Canton for most of the eighteenth century (Cheong 1997: 33; Van Dyke 2011: 79). Unsurprisingly, the designation in 1757 of Canton as the exclusive entrepôt for foreign trade spurred more Hokkiens to relocate from southern Fujian to Canton. Their number was so significant that it warranted establishing an association to promote Hokkien identity and solidarity for colleagues residing in Canton (Cheong 1997: 162). Two of the most influential *hong* merchants in the eighteenth- and nineteenth-century Canton had family origins in southern Fujian. Poankeequa or Puankhequa 潘啓官 (1714–88), the founder of the most successful international trading company of Qing China, was appointed in 1760 as the first head of the Hong Association of Canton, as advocated by nine major *hong* merchants. Howqua II 浩官 (1769–1843; full name: Wu Bingjian 伍秉鑑) was recognized as the richest merchant in Chinese history, and his family origins

---

21   See Perdue (2010) for a vivid description of this system.

can be traced to Quanzhou. The Wu's family business had evolved, over generations, from tea farming in southern Fujian to tea trading in Canton (Wang 2011, chapter 9).

Chinese trade with Southeast Asia was conducted in a rather different manner. It was Chinese merchants who played the itinerant role, travelling back and forth between China – southern Fujian, more specifically – and Southeast Asia. As no restrictions were imposed in Southeast Asia on Europeans learning Chinese or Chinese people learning foreign languages, freedom of direct interactions between Europeans and Chinese merchants naturally gave rise to significant language contact not possible under the Canton Trade System. In fact, it was in Malacca that pioneering work by Englishmen on Chinese linguistics came to fruition in the early nineteenth century, yielding Morrison's *A Grammar of the English Language* 英國文語凡例傳 (1823a) for Chinese learners at the Anglo-Chinese College, and Medhurst's *A Dictionary of the Hok-Këèn Dialect of the Chinese Language* (1832), the first bilingual dictionary for Southern Min. These endeavours shed light on the linguistic conditions of the Malay Peninsula and suggest vigorous Anglo-Chinese language contact, especially with Southern Min, prior to the formal establishment of British Malaya.

5.2   *The Emergence of Chinese Pidgin English*

According to Van Dyke (2005: 77–78), until the tightening of foreign trade in China in the mid-eighteenth century, communication between European and Chinese merchants relied greatly on the service of bilingual interpreters from Macao, who spoke Portuguese and Cantonese. By the early 1730s, however, it was reported that a simplified form of English, which would develop into what became known as Pidgin English, represented an indispensable communicative tool for foreign trade in Canton. Van Dyke (2011: 13, 124) further notes that throughout most of the eighteenth century, the term "European language" or "foreign language" was invariably understood by Sino-European traders to be Pidgin English. A licensed linguist, who had to be Chinese, was required to be fluent in Pidgin English, Cantonese, and Mandarin.

A curious comment, however, came from the pioneering Sinologist Robert Morrison (1823b): "not one of the five licensed linguists in Canton could read or write any foreign language and were not necessarily very skilled even in their own language".[22] His comment aptly reflected the sociolinguistic complexity of early-modern China, where diglossia was widespread due to the lack of a national language for verbal communication (Chen 1999). Obliquely, this

---

22   This quotation is taken from Van Dyke (2005: 78).
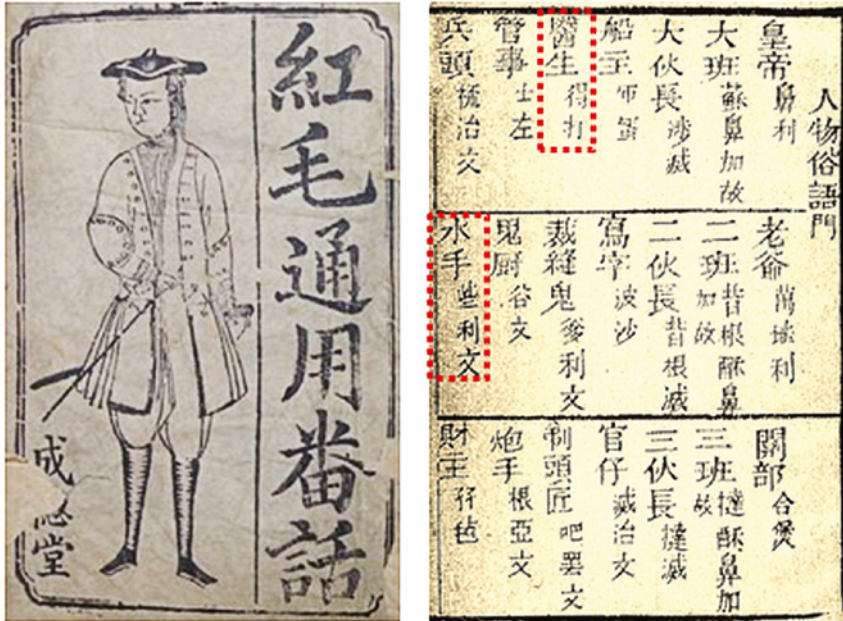
FIGURE 1.4 General foreign speech of the Europeans 紅毛通用番話, with a sample page
PHOTOGRAPHS WITHOUT COPYRIGHT INFORMATION. THE BOOK COVER
ON THE LEFT, HTTPS://VOICES.UCHICAGO.EDU/ARTPOLITICSEASTASIA/
FILES/2020/10/UNTITLED.JPG; SAMPLE PAGE ON THE RIGHT, ADAPTED
FROM HTTPS://STATIC.HKEJ.COM/HKEJ/IMAGES/2016/11/30/1444618_6
66771E68FE9B9CA5DE3C888A8A6DB0B.JPG (EMBEDDED IN THE ONLINE
ARTICLE: HTTPS://MONTHLY.HKEJ.COM/MONTHLY/ARTICLE/ID/1444618/
華人學習英語用書和首部「英漢雙語詞典」)

statement also reveals a dynamic linguistic environment of Sino-European
trade from the sixteenth century, involving Portuguese, Dutch, English, and
other European languages on the southern and south-eastern shores of China,
where a wealth of Sinitic varieties long abounded. As discussed previously, such
extensive Sino-European contact was not confined to southern China; rather,
it took place all over Southeast Asia, from the mainland to the archipelagos.

Pidgin English, by its very nature, was employed to enable the otherwise
impossible exchange of messages between speakers of different languages;
literacy in this variety was not expected. Early glossaries of Chinese Pidgin
English, as exemplified in Fig. 1.4, often utilized Chinese characters to approxi-
mate the pronunciation of English words without providing their actual
spelling in the Latin alphabet, e.g. 得打 (Cantonese: *dak daa*) for *doctor*,
些利文 (Cantonese: *se lei man*) for *sailorman*, and 痕甚 (Cantonese: *han sam*)

for *handsome*, etc.[23] Using this approach, the "linguists" in Canton naturally would not have been able to read or write any European language. What is remarkable about Morrison's comment lies in the second part, concerning their seemingly insufficient command of Cantonese and/or Mandarin (judged by his mastery of Mandarin and Cantonese). This suggests that some of these interpreters may have spoken a Sinitic language other than Cantonese and Mandarin as their mother tongue. If this had been the case, the most likely candidate would have been Southern Min.[24]

The Pidgin English names of influential Hokkien tycoons, as seen in the records of European traders, reflected their Southern Min pronunciation (Cheong 1997: 24). A name, initially the personal name of the founder of a trading company, would be passed on to successors who inherited the family business, followed by a Roman numeral. Table 1.2 displays names of some of the most famous *hong* leaders, transcribed in the three Chinese languages most important to early Sino-European trade. At the top of the table, the merchant names are provided in Chinese characters, followed by the year when their name was first recorded in foreign trade and the hometown or family origins of the merchant.[25] The rhyming patterns of these Chinese varieties confirm the Hokkien identity of these guild leaders. The absence of a word-final /n/ in the English spelling of the Chinese honorific suffix *qua* 官 is particularly revealing, as a complete loss of nasality in the rhyme is possible only in certain dialects of Southern Min such as Burmese Hokkien (e.g. 三 'three'; Mandarin: *san*, Cantonese: *saam*, Burmese Hokkien: *sa*).[26] This evinces frequent contact and interaction between European merchants and their Southern Min-speaking counterparts in the early period of Sino-British trade.

---

23 See Li et al. (2005) and Qiu (2017) for further discussions on self-taught materials for learning Chinese Pidgin English. This practice of pidgin learning appears to have been the norm in East Asia, where the writing system of local languages differs significantly from that of European languages. Further examples can be found in Atkinson's (1879) exercises of pidgin Japanese for English speakers; Japanese words were represented in English words without any *kana* or *kanji*, e.g. *nanny* for *nani* 何 'what', *yachts* for *yatsu* 八 'eight', and *oh my* for *omae* お前 '(singular) you'.

24 Although Van Dyke (2005) devoted a chapter to detail duties and income of interpreters working in the foreign trade, unfortunately, personal background of these Chinese interpreters was not available.

25 Sources of detail are Van Dyke (2011) and Cantonese Wikipedia (https://zh-yue.wikipedia .org).

26 The weakening of final nasals has resulted in varying effects in Southern Min dialects. In some dialects, the word-final /n/ was lost with a compensatory nasal vowel on the rhyme (similar to what has happened historically in French). In others, deletion of /n/ is not accompanied by nasalization, but simply by a removal of the coda.

TABLE 1.2   Transcriptions of the name of famous *hong* 行 merchants in Canton

	陳壽官; 1716 Fujian	潘啓官; 1730s Quanzhou	浩官; 1784 Quanzhou	鼇官; 1811 Zhangzhou
English	Tan Suqua	Poankeequa I	Howqua I	Goqua I
Hokkien	tan su kwa	phwan khey kwa	hao kwa	(n)go kwa
Cantonese	chan sau gun	pun kai gun	hou gun	ngou gun
Mandarin	chen shou guan	pan ji guan	hao guan	ao guan

While the personal background of Chinese interpreters is generally lacking in studies on early Sino-European trade, comments on the linguistic ability of Chinese commercial leaders can occasionally be found. For instance, Poankeequa was documented to have travelled on junks from his home village in southern Fujian to Manila at a young age and developed skills in Spanish and (Pidgin) English (Perdue 2010; Van Dyke 2016: 61).[27] Cai Hunqua, whose Chinese name is identified as 蔡煌官 in Van Dyke (2016: 3), was another merchant with whom Europeans could communicate directly in Pidgin English (Van Dyke 2011: 127). Excluding the honorific suffix *qua*, the pronunciation of his name is as follows: *cai huang* in Mandarin, *choy wong* in Cantonese, and *chai hong* (or *chua hong*) in Southern Min. The Cantonese pronunciation is the least similar to the English spelling. In Mandarin and Southern Min, the initial consonants of the individual characters are identical despite their varying romanizations. Therefore, the focus lies in the rhyme of the characters. In Southern Min, the family name becomes homophonous to its Mandarin equivalent when it is read as *chai*, which also gives rise to the English spelling. While neither Mandarin nor Southern Min shares the precise rhyme of his first name as seen in the English spelling, the Southern Min pronunciation is much closer than the Mandarin one. This indicates a Hokkien identity of Cai Hunqua. Both Hokkien merchants were actively engaged in Sino-European trade since the 1730s or shortly afterwards and could be regarded as the earliest group of intermediaries who conversed with Europeans in Chinese Pidgin English.

5.3   *Southern Min as a Donor to Chinese Pidgin English*
A prolonged, typically peaceful contact between speakers of two languages in a more or less equal relation may give rise to a new variety with mixed features

---

27   Poankeequa's hometown was originally under the administration of Quanzhou, but was merged into Zhangzhou in 1958.

derived from both languages. One example is Pu-Xian 莆仙, the result of profound contact between speakers of Southern Min and Eastern Min. Language contact as a side product of European imperialism, by contrast, is characterized by marked power differences between European and local speakers. The resultant pidgin and creole languages are typically formed with a hegemonic language donating its vocabulary and a marginalized language providing the grammatical backbone. Accordingly, Chinese Pidgin English predictably demonstrates a Chinese grammatical structure on an English lexical foundation. Yet reality is more complicated in terms of its vocabulary. As discussed above, an extremely diverse linguistic setting characterized the trade between European merchants and Chinese middlemen from the late 1500s to the 1730s, before the emergence of Chinese Pidgin English. Sino-European commerce can in fact be considered an addition to well-established networks of maritime trade in the Far East. Under such circumstances, Chinese Pidgin English received lexical contributions from a wealth of donor languages. Table 1.3 provides a selection of words donated by languages other than English, based on two short glossaries compiled by Leland (1876) and Airey (1906).[28] To facilitate the recognition of expressions derived from Sinitic sources, Chinese characters are supplemented to the etyma in the right column.

TABLE 1.3　Donor sources of common vocabulary in Chinese Pidgin English

Word	English gloss	Etymon
compladore ~ kam-pat-to	comprador; steward	Portuguese *comprador*
cumshaw	a present; a tip, gratuity	Hokkien *kam sia* 感謝
fa ts'ai	to get rich	Mandarin *fa cai* 發財
fytie	go quick; hurry!	Cantonese *fai di* 快啲
hahng	great firms which formerly regulated all Chinese commerce	Hokkien *hang* 行 (cf. Mandarin *hang*)
hong	great firms which formerly regulated all Chinese commerce	Cantonese *hong* 行
hwan-na-kou	dog of European breed	Hokkien *hwan a kow* 番仔狗
jin-rick-sha	a vehicle like a Bath chair	Japanese *jinrikisha*
kung-he	congratulations	Cantonese *gung hei* 恭喜

28　For the Chinese vocabulary in Pidgin English, Leland (1876) specified only Mandarin and Cantonese sources, while Airey (1906) mentioned Cantonese and Mandarin as the source for a few items.

TABLE 1.3   Donor sources of common vocabulary in Chinese Pidgin English (*cont.*)

Word	English gloss	Etymon
ming-pak	to understand clearly	Cantonese *ming baak* 明白
pylong	thief; to thieve	Hokkien *phai lang* 歹人
sa-ki	a Japanese alcoholic drink	Japanese *sake*
savvy ~ sha-pi	to know; to understand	Portuguese *saber*
taotai	a Chinese magistrate	Mandarin *dao tai* 道台

As shown previously in Fig. 1.4, the Chinese made use of Chinese characters to indicate pronunciations of European words. This method resulted in alternate spellings of pidgin words based on the sounds of Chinese, as seen in *compladore* vs. *kam-pat-to* and *savvy* vs. *sha-pi*. Due to the extensive trade networks in the Far East, even the Japanese contributed a few culture-related words to the repertoire of Chinese Pidgin English.

Among the Sinitic languages, Cantonese donated the largest number of words to Pidgin English, while Hokkien and Mandarin also contributed their fair share. Unlike the character-based diffusion of *ang moh* 紅毛 and the tropical fruit names discussed above, Sinitic vocabulary typically entered Pidgin English through oral rather than written communication. The provenance of its vocabulary was irrelevant to the users of this improvised communicative tool. Some words exhibit doublets, as is the case for 行, which was the official firm in charge of handling Sino-European trade: Cantonese *hong* versus Southern Min and Mandarin *hahng*.

The honorific suffix *qua* 官 and the loanword *cumshaw* 感謝 are probably among the earliest lexical contributions from Hokkien to Pidgin English. The former was used not only after names of Hokkien merchants, but also for other well-respected Chinese, e.g. Mowqua 茂官 and Kingqua 經官 – both were famous Cantonese merchants. Remarkably, the meaning of *cumshaw* 感謝 has changed from 'to thank' to 'a present; a tip, gratuity', denoting a common practice of bribery in order to smoothen things in trading with the Chinese. This usage as a noun is not attested in Southern Min. Finally, it is worth noting that *The American Cyclopædia: A Popular Dictionary of General Knowledge*, under the presumption of English being the supplier of vocabulary to Pidgin English, treated *commission* as the etymon of *cumshaw* 感謝 with a change of meaning from 'compensation for services' to 'gratuity' (Ripley & Dana 1875:

507). By contrast, the Southern Min origin of this term was readily perceived by contemporary Dutch Sinologists (Kuiper 2017: 502).

5.4   Phi jun 批准: From 'Approve' to 'Business'

The primary duty of an interpreter was to convey messages as accurately as possible. However, the "linguists" servicing foreign and Chinese merchants in Canton had other important duties. Van Dyke (2005: 79–80) reports that "bills of lading, stamped with the appropriate chops, had to accompany all merchandise and bullion shipped between Whampoa and Canton, and it was the job of the linguists to obtain these documents". It is not difficult to imagine numerous permits, which had to be obtained before shipments could be unloaded ashore and would complete the international trade process. From the perspective of European merchants, official approval to unload all cargos from a ship was the final step to a successful long-haul voyage from Europe to China, which would take months depending on the season. No permits effectively meant no business.

Considering the importance attached to official permits in Sino-European trading in Canton, I hypothesize that the frequent use of 批准 (Mandarin: *pi zhun*, Southern Min: *phi jun* or *phwey jun*) 'to approve' has given rise to a semantic change from 'to approve' to 'business' when the term entered Chinese Pidgin English. This is reminiscent of Southern Min *cumshaw* 感謝, which has changed from 'to thank' to 'a present; a tip, gratuity' in Chinese Pidgin English. The semantic shift of *phi jun* 批准 could be explained in terms of re-analysis, where the intended message of "(cargos have been) approved" was interpreted by foreign merchants as something along the lines of "(successful) business".

The term 批准 has variant pronunciations across Southern Min dialects:[29] *phi jun* or *phey jun* in Zhangzhou and *phwey jun* in Xiamen and Quanzhou (also compare *pai zeon* in Cantonese and *pi zhun* in Mandarin). Phonologically, the case for *phi jun* 批准 as the origin of *pidgin* is appealing, as its pronunciations in Zhangzhou Hokkien and Mandarin demonstrate a striking similarity to English *pigeon*, the earliest attested spelling for *pidgin*. Yet treating Mandarin as the donor language for *pidgin* poses difficulties regarding the observed change from verb to noun. While this is common in Japanese, which systematically converts borrowed verbs into nouns,[30] such a practice requires

---

29   For details, see http://www.zdic.net/zd/yy/my/批.

30   For example, *anaunsu* アナウンス 'announcement' (from English *announce*). Japanese borrows many Chinese verbs, such as *xu nuo* 許諾 'to promise', *po chan* 破產 'to bankrupt',

a morphological motivation. In Sinitic languages, *phi jun* 批准 functions as a verb rather than a noun. Unless it is regarded as an exceptional case, as in *cumshaw* 感謝 mentioned previously, the borrowing of a verb from Chinese into Pidgin English as a noun, which simultaneously underwent a semantic change, would be rather extraordinary. Prior to this lexical borrowing, some linguistic process must have taken place to pave the way for the change of *phi jun* 批准 from 'to approve' to 'permit' and then to 'business'.

Chinese terminology in the domain of commerce and/or laws contains such terms as *pi zhun wen jian* 批准文件 'permit' (lit. 'approve document') and *jin kou pi wen* 進口批文 'import permit' (lit. 'import approve document'). The former is a nominal compound with a verb-plus-noun structure, whereas the latter represents a similar structure based on a shortened form of the former; see (5a). This kind of clipping, the shortening of quadrisyllablic nouns to disyllables, is extremely productive in Mandarin. Such quadrisyllabic compounds are nouns consisting of two disyllabic parts. Mandarin employs a selective clipping strategy to shorten each constituent by one syllable. A total of four patterns are possible, with decreasing productivity, as exemplified from (5a) to (5d). In most cases, the first syllable of each formative is retained, as shown in (5a).

(5a)   Selective clipping on the second and fourth syllables
       pi zhun wen jian 批准文件 'approval document; permit' → *pi wen*
       ding qi cun kuan 定期存款 'fixed-term deposit' → *ding cun*
       bei jing da xue 北京大學 'Peking University' → *bei da*

(5b)   Selective clipping on the second and third syllables
       nong ye yin hang 農業銀行 'Bank of Agriculture' → *nong hang*
       tai wan tong bao 臺灣同胞 'compatriot of Taiwan' → *tai bao*

(5c)   Selective clipping on the first and third syllables
       xiang gang tong bao 香港同胞 'compatriot of Hong Kong' → *gang bao*
       xiang gang di qu 香港地區 'district of Hong Kong' → *gang qu*

(5d)   Selective clipping on the first and fourth syllables
       xiang gang da xue 香港大學 'University of Hong Kong' → *gang da*
       shen ti jian cha 身體檢查 'check-up of body' → *ti jian*

---

yin tui 引退 'to recede' etc., and converts them into nouns. To use these as verbs, they must appear in the construction 'N + *suru*' (i.e. 'to do N') so that the Japanese verbal morphology can be applied through the native verb *suru*.

Tail-clipping represents another manner to shorten nouns from quadri-syllablic to disyllabic in Mandarin. In terms of style and register, it is casual and informal, and requires support from discursive or linguistic contexts. For instance, the truncated form *huo qi* 'general deposit' in (6) is normally employed in such compounds as *huo qi hu kou* 活期戶口 'general deposit account'. Tail-clipping is seldom observed in Standard Chinese.

(6)　Tail-clipping in Mandarin
　　　huo qi cun kuan 活期存款 'general deposit' → *huo qi*
　　　qing hua da xue 清華大學 'Peking University' → *qing hua*

All shortened forms presented in (5) can be used in Cantonese, often with a tint of borrowing. In colloquial Cantonese, however, the preferred clipping strategy targets the second constituent as a whole ("tail-clipping"), rather than individual syllables in both constituents ("selective clipping"). As shown in (6) and (7), tail-clipping curtails a longer noun by deleting the entire second con-stituent. This means that the shortened forms inevitably suffer from the partial loss of lexical content. This explains the necessity of a high degree of support from the discursive context. Note that the part-of-speech of a shortened form always remains intact, and this lends great service to coping with ambiguity which may arise on the surface. For instance, *ding kei* 定期 'regularly' is an existent word in Cantonese. The fact that it is an adverb facilitates its disam-biguation from the shortened form *ding kei* 定期 'fixed-term deposit', which is a noun. Moreover, (8) shows that selective clipping on the syllable level also constitutes an important strategy for generating abbreviations in Cantonese.

(7)　Tail-clipping in Cantonese
　　　ding kei cyun fun 定期存款 'fixed-term deposit' → *ding kei*
　　　coeng tou din waa 長途電話 'long-distance call' → *coeng tou*
　　　sau tai hang lei 手提行李 'hand-carried luggage' → *sau tai*
　　　dung aa ngan hong 東亞銀行 'East Asia Bank' → *dung aa*
　　　hang sang ngan hong 恆生銀行 'Hang Seng Bank' → *hang sang*

(8)　Selective clipping in Cantonese
　　　pai zeon man gin 批准文件 'approval document; permit' → *pai man*
　　　hang sang zi sou 恆生指數 'Hang Seng index' → *hang zi*
　　　gung gung fong uk 公共房屋 'public housing' → *gung uk*
　　　ning mung ho lok 檸檬可樂 'coke served with fresh lemon' → *ning lok*
　　　gai daan saam man zi 雞蛋三文治 'egg sandwich' → *daan zi*

Both Southern Min and Cantonese appreciate simplicity of tail-truncation in spite of the potential ambiguity which may result from this type of truncation.[31] Southern Min differs from Cantonese, as well as Mandarin, in its tendency to refrain from selective clipping.[32] Various examples of tail-clipping in Southern Min are presented below.

(9)   Productive tail-clipping of nouns in Southern Min
      ting ki chun khwan 定期存款 'fixed-term deposit' → *ting ki*
      tng toh tien wey 長途電話 'long-distance call' → *tng toh*
      kia oh i in 鏡湖醫院 'Kiang Wu Hospital' → *kia oh*[33]
      swa ting i in 山頂醫院 'St. Januario Hospital' → *swa ting*

In addition to abbreviating quadrisyllabic nouns to disyllabic ones as exemplified in (9), Southern Min can also shorten a verb phrase to a disyllabic expression, as seen in (10). This type of shortening is unknown in Mandarin and Cantonese. Even in Southern Min, it is much more restricted than the clipping of quadrisyllabic nouns. Tail-clipped verb phrases tend to be sensitive to lexical collocation instead of linguistic factors; the clipping becomes unacceptable when the head of the verb phrase is replaced, as shown in (10').

(10)   Restricted tail-clipping of verb phrases in Southern Min
       jia hun ki 食煙枝 'to smoke (cigarette)' → *jia hun*
       jia ping tiao 食冰條 'to eat a popsicle' → *jia ping*
       sio kim un jwa 燒金銀紙 'to burn joss paper' → *sio kim*

(10') *wey hun ki* 買煙枝 'to buy (cigarette)' → ? *wey hun* (acceptable in
       Taiwanese)
       wey ping tiao 買冰條 'to buy a popsicle' → *wey ping* (lit. 'to buy ice')
       wey kim un jwa 買金銀紙 'to buy joss paper' → *wey kim* (lit. 'to buy gold')

Finally, the quadrisyllabic nouns in (11) share an identical structure of verb-plus-noun. In colloquial Hokkien, the use of *kia kwa ho* 'to send by registered mail' in lieu of *kia kwa ho iu kia* 寄掛號郵件 is quite acceptable. While the

---

31   Tail-truncation is also applicable to loanwords in Cantonese, e.g. *inso* 'insurance' and *sitkiu* 'security guard'. This kind of truncation is also sometimes found in Japanese, generating disyllabic loanwords based on the initial part of the English original, e.g. *kiro* キロ from *kilometer* or *kilogram* and *suupa* スーパ from *supermarket*.

32   See also Churchman (this volume) on related processes in Penang Hokkien.

33   The names of 鏡湖醫院 and 山頂醫院, the two major hospitals in Macao, are often tail-truncated by the locals in Cantonese and Southern Min.

shortened form *phi jun* 批准 'permit' is not observed in present-day Southern Min (probably due to change of society and loss of its daily use context), it is theoretically a possible abbreviation produced by tail-clipping.

(11)   Possible tail-clipping of nouns in Southern Min
       kwa ho iu kia 掛號郵件 'registered mail' → *kwa ho*
       phi jun mun kia 批准文件 'approval document; permit' → *phi jun*

Returning now to the etymological path from Southern Min *phi jun* 批准 to English *pidgin*, the word in question may thus be argued to be a shortened form of *phi jun mun kia* 批准文件 'permit', whose meaning was re-analysed as '(permit for) business'. Considering the variant pronunciations of 批准 (*phi jun* or *phey jun* in Zhangzhou, *phwey jun* in Xiamen) in Southern Min, the Hokkiens who introduced this word into Pidgin English were probably natives of Zhangzhou. Alternatively, one might conjecture that a Hokkien middleman simply code-switched to Mandarin *pi zhun* 批准 in his conversation with Europeans. In any event, the emergence of *pidgin* involves a series of processes, as presented in Fig. 1.5.

   This hypothesis also accounts for the variant pronunciations of *pidgin* 'business' found in self-taught materials of Pidgin English for Cantonese speakers, i.e. *bit zin* 必剪 versus *bei zin* 卑剪. Such a variation corresponds to the dialectal variation of 批准 (*phi jun, phey jun, phwey jun*) in Southern Min. Due to Cantonese phonotactic constraints, the labiovelar glide in the third dialectal form is subject to deletion. Therefore, Cantonese speakers may have borrowed the term 批准 (*phi jun, phey jun*) 'business' directly from Southern Min in their speech of Pidgin English. Alternatively, an oblique borrowing from Pidgin English, as indicated by the broken line in Fig. 1.5, is equally plausible.

6   Concluding Remarks

Against the backdrop of the sociohistorical background of Southern Min between the sixteenth and the nineteenth centuries, this chapter has discussed some important contributions made by this less-studied Sinitic language in Sino-European contact, which first took place in Southeast Asia and then in coastal China. Terms such as *ang moh* 紅毛 'European, white person', *han ji* 番薯 'sweet potato', *toh lien* 榴槤 'durian', and *ang moh tan* 紅毛丹 'rambutan' all trace their origin to Southern Min. The earliest written record of *ang moh* 紅毛 is attested in a text of early seventeenth century, in which the Dutch, already known by the locals as *ang moh* 紅毛, made their first visit to Macao in

Southern Min   *phi jun mun kia* or *ph(w)ey jun mun kia*

'(business) permit'

① tail truncation

phi jun   or   *ph(w)ey jun* '(business) permit'

② semantic re-analysis

phi jun   or   *ph(w)ey jun* 'business'

③ code-switching
to Mandarin *pi zhun*

Pidgin English   *pigeon* 'business' (*pidgin*, later)

re-borrowing

Pidgin English
by Cantonese   *bit zin* (必剪) 'business'   or   *bei zin* (卑剪)
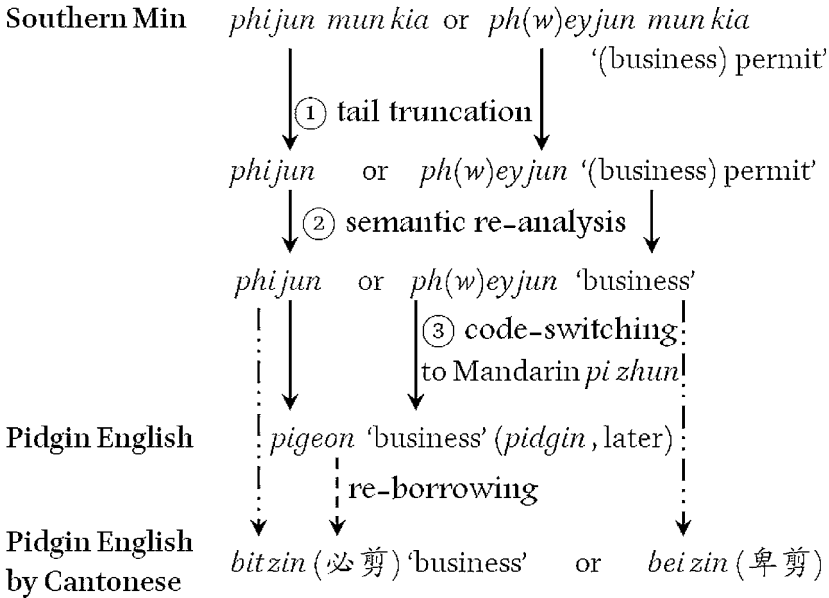
FIGURE 1.5   Linguistic processes in deriving pidgin from Southern Min *phi jun* or
ph(w)ey jun
OWN WORK

1601. This account suggests that this epithet was created in the late 1590s, after the Hokkien in Indonesia first encountered the Dutch. The borrowing of *toh lien* 榴槤 'durian' and *ang moh tan* 紅毛丹 'rambutan' from Malay into generic Chinese through Southern Min witnesses the extensive contact between Hokkiens and the people of Maritime Southeast Asia. By virtue of such a relationship, the sweet potato was introduced into China by a native of Hokkien from the Philippines, rendering Southern Min the language in which the naming of this crop first took place.

Since the early eighteenth century, Hokkien merchants had played a crucial role in international trade in Canton. Many business leaders were descendants from southern Fujian, and some of them were reported to communicate directly with European merchants in Pidgin English. The Southern Min word *phi jun* or *ph(w)ey jun* 批准 'to approve', then, is arguably the source of English *pidgin*. A series of linguistic processes took place prior to the introduction of this Southern Min word into Pidgin English. Firstly, *phi jun mun kia* or *ph(w)ey jun mun kia* 批准文件 'approval document; permit' was shortened to *phi jun* or *ph(w)ey jun* through tail-clipping – a morphological process common in Southern Min and Cantonese, but relatively rare in Mandarin – and

then its meaning was re-analysed from '(business) permit' to 'business'. Code-switching involving Mandarin *pi zhun* 批准 may have occurred before the variant *ph(w)ey jun* made its way into Pidgin English. The dialectal variation in Southern Min – *phi jun* and *ph(w)ey jun* – is probably responsible for the varying forms *bit zin* 必剪 and *bei zin* 卑剪 for *pidgin* 'business' as found in the Pidgin English of Cantonese speakers.

From *ang moh* 紅毛 to *phi jun* 批准, the timeline in Table 1.4 indicates significant historical events pertaining to the linguistic issues investigated in this chapter. In the case of *ang moh* 紅毛, the estimated time of its first occurrence is confined to a short interval. Conversely, the estimated timeframe for the first appearance of *phi jun* 批准 'business' in Chinese Pidgin English spans decades, most likely taking place around the mid-eighteenth century. Cumulatively, these findings from Southern Min vocabulary support a central role of Hokkien speakers, including those based in Southeast Asia, in the history of Sino-European contact and the linguistic expressions created in this process.

TABLE 1.4   Timeline of relevant historical events and linguistic matters

By 1410s	Hokkien community formed in Palembang
By 1420s	Hokkien community formed in Sumatra
1567	Zhangzhou as China's sole seaport for foreign trade
1594	China's large scale planting of sweet potatoes initiated in Fujian
1596	The first Dutch arrival in Java
...	*ang moh* 紅毛 used by Hokkiens to refer to the Dutch
1601	The first Dutch visit to Macao
1600s	The first written record of 紅毛 attested in *Yue Jian Bian* 粵劍編
...	
1685	Canton (in addition to Zhangzhou) opened for international trade
1700s	Foreign trade in Canton dominated by Hokkien merchants
By 1730s	The Southern Min honorific suffix *-qua* 官 used in English
1730s	Precursor of Chinese Pidgin English reported in Canton
...	Southern Min *phi jun* 批准 '(business) permit' re-analysed as *pigeon* 'business' in Chinese Pidgin English
1757	Canton as China's sole entrepôt ("Canton Trade System")
1760	Poankeequa 潘啓官 as the first head of *Hong* 行 Association of Canton
1800s	Booklets for self-taught Pidgin English printed in Canton
1842	Sino-British Treaty of Nanking (to open Canton, Amoy, Fuzhou, Ningbo & Shanghai for trade)
1859	The spelling of *pigeon* 'business' first recorded in English

Acknowledgements

References

Airey, Frederick, *Pidgin Inglis Tails and Others* (Shanghai: Kelly & Walsh, 1906).

Akamine, Mamoru, *The Ryukyu Kingdom: Cornerstone of East Asia* (ed. Robert Huey, transl. Lina Terrell, Honolulu: University of Hawai'i Press, 2017).

Andrade, Tonio, *How Taiwan Became Chinese: Dutch, Spanish, and Han Colonization in the Seventeenth Century* (New York: Columbia University Press, 2007).

Ansaldo, Umberto, Stephen Matthews and Geoff P. Smith, 'China Coast Pidgin: Texts and contexts', *Journal of Pidgin and Creole Languages*, 25 (2010): 63–94.

Arai, Hakuseki 新井白石, *Narratives of the Occident* (西洋紀聞 Seiyō Kibun), 1715. Available at http://ja.wikisource.org/wiki/西洋紀聞 (Retrieved on March 15, 2020).

Atkinson, H., *Revised and Enlarged Edition of Exercises in the Yokohama Dialect* (Yokohama: Japan Gazette Office, 1879).

Chen, Ping, *Modern Chinese: History and sociolinguistics* (Cambridge: Cambridge University Press, 1999).

Cheng, Lim-keak, *Social Change and the Chinese in Singapore: A socio-economic geography with special reference to bāng structure* (Singapore: Singapore University Press, 1985).

Cheong, Weng Eang, *The Hong Merchants of Canton: Chinese merchants in Sino-Western trade* (Richmond, Surrey: Curzon Press, 1997).

De Witt, Dennis, *History of the Dutch in Malaysia* (Petaling Jaya, Malaysia: Nutmeg Publishing, 2009).

Ding, Picus Sizhi, *Southern Min (Hokkien) as a Migrating Language: A Comparative Study of Language Shift and Maintenance across National Borders* (Singapore: Springer, 2016).

Downs, Jacques M. and Frederic D. Grant Jr., *The Golden Ghetto: The American Commercial Community at Canton and the Shaping of American China Policy, 1784–1844* (new edn., Hong Kong: Hong Kong University Press, 2014).

Goodman, Grant K., *Japan and the Dutch 1600–1853* (Richmond, Surrey: Curzon, 2000).

Ho, Ping-ti, 'The introduction of American food plants to China', *American Anthropologist*, 57/2 (1955): 191–201.

Holm, John, *An Introduction to Pidgins and Creoles* (Cambridge: Cambridge University Press, 2000).

Kshetry, Gopal, *Foreigners in Japan: A Historical Perspective* (Bloomington, IN: Xlibris Corporation, 2008).

Kuiper, Koos, *The Early Dutch Sinologists (1854–1900): Training in Holland and China, Functions in the Netherlands Indies* (Leiden: Brill, 2017).

Lach, Donald F., *Asia in the Making of Europe, Volume I: The Century of Discovery* (new edn., Chicago: University of Chicago Press, 1994).

Lee, Khoon Choy, *Golden Dragon and Purple Phoenix: The Chinese and their Multi-Ethnic Descendants in Southeast Asia* (Singapore: World Scientific, 2013).

Leland, Charles G., *Pidgin-English Sing-Song, or, Songs and stories in the China-English dialect. With a vocabulary* (London: Paul, Trench, Trübner & Co., 1876).

Levathes, Louise, *When China Ruled the Seas: The Treasure Fleet of the Dragon Throne, 1405–1433* (New York: Oxford University Press, 1996).

Li, Longsheng 李隆生, *International Trade of the Qing Dynasty* 清代的國際貿易 (Taipei: Showwe, 2010).

Li, Michelle, Stephen Matthews and Geoff P. Smith, 'Pidgin English texts from the Chinese English Instructor', in Geoff P. Smith and Stephen Matthews (eds.), *Chinese Pidgin English: Texts and contexts*. Special issue of the *Hong Kong Journal of Applied Linguistics*, 10 (2005): 79–167.

Li, Rulong 李如龍 and R. Yao 姚榮松 (eds.), *The Southern Min Topolect* 閩南方言 (Fuzhou: Fujian People Press, 2008).

Ma, Huan 馬歡, *The Overall Survey of the Ocean Shores* 瀛涯勝覽 (1451). Available at https://ctext.org/library.pl?if=gb&res=105804 (Retrieved on March 5, 2020).

Medhurst, Walter Henry, *A Dictionary of the Hok-Këèn Dialect of the Chinese Language* (Macao: East India Company's Press, 1832).

Miksic, John, *Singapore and the Silk Road of the Sea 1300–1800* (Singapore: NUS Press, 2013).

Morrison, Robert, *A Grammar of the English Language: For the use of the Anglo-Chinese College* (Macao: East India Company's Press, 1823a).

Morrison, Robert, *Notices Concerning China and the Port of Canton* (China: Mission Press, 1823b).

Ng, Chin-keong, *Trade and Society: The Amoy Network on the China Coast 1683–1735* (second edn., Singapore: NUS Press, 2015).

Ni, Zhijia 倪志佳 and Xiang Mengbing 项梦冰, 'Chinese word geography: A case study of *fanshu* "sweet potato"' 番薯的词汇地理学研究, *Modern Linguistics* 现代语言学, 2/1 (2014): 44–67. http://dx.doi.org/10.12677/ML.2014.21007.

Norman, Jerry, *Chinese* (Cambridge: Cambridge University Press, 1988).

Perdue, Peter C., 'The rise & fall of the Canton Trade System – I', (2010). Downloadable at https://visualizingcultures.mit.edu/rise_fall_canton_01/pdf/cw_essay.pdf (Retrieved on July 23, 2018).

Qiu, Zhihong 邱志红, '"Devil talk" in the East: A study of early English glossaries of the type "foreign speech of the redhead"' "鬼话"东来: "红毛番话"类早期英语词汇书考析, *Study on the Qing History* 清史研究, 2 (2017): 113–21.

Ripley, George and Charles Anderson Dana (eds.), *The American Cyclopædia: A Popular Dictionary of General Knowledge* (volume 13, New York: Appleton, 1875).

Simoons, Frederick J., *Food in China: a Cultural and Historical Inquiry* (Boca Raton, FL: CRC Press, 1991).

Song, Ong Siang, *One Hundred Years' History of the Chinese in Singapore* (Singapore: University of Malaya Press, 1967).

Tamura, Eileen H. and Linda Menton, *China: Understanding its Past* (Hololulu: University of Hawaii Press, 1997).

Tang, Wenji 唐文基 (ed.), *An Economic History of Ancient Fujian* 福建古代經濟史 (Fuzhou: Fujian Education Publishing, 1995).

Twitchett, Denis C. and Frederick W. Mote (eds.), *The Cambridge History of China, Volume 8, Part 2: The Ming Dynasty, 1368–1644* (Cambridge: Cambridge University Press, 1998).

Uchida, Keiichi 内田慶市, 'Pidgin – 異語言文化接觸中的一種現象' / 'Pidgin English as a phenomenon of language and cultural interactions', *Journal of East Asian Cultural Interaction Studies* 東アジア文化交渉研究 (関西大学文化交渉学教育研究拠点), 2 (2009): 197–207.

van der Loon, Piet, 'The literature of the regional languages of Southeast China', in Annamaria Baldussi (ed.), *Le Minoranze e le Autonomie Regionali in Asia e in Africa Raffronto con il Caso Sardegna* (Orientalia Karalitana: Quaderni dell'Istituto di Studi Africani e Orientali, Numero Speciale, 1991): 89–94.

Van Dyke, Paul A., *The Canton Trade: Life and Enterprise on the China Coast, 1700–1845* (Hong Kong: Hong Kong University Press, 2005).

Van Dyke, Paul A., *Merchants of Canton and Macao: Politics and Strategies in Eighteenth-Century Chinese Trade* (Hong Kong: Hong Kong University Press, 2011).

Van Dyke, Paul A., *Merchants of Canton and Macao: Success and Failure in Eighteenth-Century Chinese Trade* (Hong Kong: Hong Kong University Press, 2016).

Walker, Hugh Dyson, *East Asia: A New History* (Bloomington, IN: Author House, 2012).

Wang, Linheng 王臨亨, *Yue Jian Bian* 粵劍編, (no date). Available at http://ctext.org/wiki.pl?if=gb&res=187407 (Retrieved on March 23, 2020).

Wang, Yuxian 王俞现, *Six Hundred Years of Merchant Groups in China* 中國商幫600年 (Beijing: CITIC Press, 2011).

Wurm, Stephen Adolphe, Rong Li, Theo Baumann, Mei W. Lee (eds.), *Language Atlas of China* (Hong Kong: Longman, 1988).

Xu, Xiaowang 徐曉望, *Study on the History of Belief in Mazu* 媽祖信仰史研究 (Fuzhou: Haifeng Publishing, 2007).

Yu, Yonghe 郁永河, *A Diary of Sulfur Mining* 採硫日記/裨海紀遊, (no date). Available at http://www.guoxue123.com/tw/01/044/008.htm (Retrieved on August 3, 2018).

Zhang, L., Q. Wang, Q. Liu and Q. Wang, 'Sweet potato in China', in G. Loebenstein and G. Thottappilly (eds.), *The Sweetpotato* (Dordrecht: Springer, 2009): 325–58.

At the Periphery of Nanyang:
The Hakka Community of Timor-Leste

Juliette Huber

1   Introduction

A story by an East Timorese elder, recounting the origin of his clan, includes
the following lines:

> My ancestor came by sea on a ship. A Chinese captain brought him
> here. My ancestor came to Irabere on a Chinese ship.

I recorded this story in 2013 among the Makalero people, who live in the remote
Iliomar subdistrict on the south-eastern coast of Timor-Leste. Clan origin sto-
ries are of great ritual significance for the Makalero, and the fact that a Chinese
captain is featured in this narrative illustrates the degree to which the ethnic
Chinese are an established part of the East Timorese scene. They started to set-
tle on the island during the Portuguese colonial period, and by 1975 a vibrant
community existed. Today, the majority of Chinese-Timorese live abroad, hav-
ing fled East Timor during the Indonesian occupation between 1975 and 1999.[1]

Despite its long history, the Chinese community of Timor-Leste has largely
been neglected in the broader literature on Chinese-descended communi-
ties worldwide. The scholarship on Chinese communities in Southeast Asia
routinely discusses communities in Burma (Myanmar), Cambodia, Indonesia,
Laos, Malaysia, the Philippines, Singapore, Thailand, Vietnam and Brunei,

---

1   In this contribution, the term "Chinese-Timorese" is used to refer to the ethnic Chinese popu-
lation of East Timor only. The ethnic Chinese population of Indonesian West Timor is not
discussed. In Timor-Leste, Chinese-Timorese are referred to, and refer to themselves, as *Cina
Timor* in Tetun, the local lingua franca (mostly using the Indonesian spelling *Cina* rather
than Tetun *Xina* in writing), or *Comunidade Chinesa Timorense* in Portuguese (cf. Kammen &
Chen 2019: 165, fn 1). Furthermore, I use "East Timor" to refer to the territory occupied today
by Timor-Leste without reference to a particular historical period. It thus covers colonial-era
Portuguese Timor, the Indonesian province of East Timor (Timor Timur; 1975–99) and inde-
pendent Timor-Leste. The nomenclature used in the discussion of specific historical periods
may differ.

but makes no mention of Timor-Leste.[2] In recent years, however, this small community has garnered somewhat more attention, having been the focus of a small number of publications aimed at academic audiences (Kammen & Chen 2019; Soares 2019; Gunn 2016; Berlie 2015; Pinto 2015; Pinto 2014; Chew & Huang 2014) on the one hand, and the general public on the other (Mulyanto 2019; Raynor 2019; England 2001; England 1999).

The aims of this chapter are twofold. As part of the present collection, it intends to increase the visibility of the Chinese-Timorese community in the broader literature on Chinese-descended people worldwide. Secondly, it will give a first preliminary account of the endangered Hakka variety spoken by that community, thus contributing to our knowledge of the properties of Hakka varieties in Southeast Asia. The chapter is organized as follows: section 2 provides background on Hakka, East Timor, and the historical links between Timor and Macao, the point of departure for many Chinese immigrants to East Timor. Section 3 outlines the history of the Chinese-Timorese community from the colonial period to the present day. Section 4 discusses the Hakka variety of Timor-Leste, drawing attention to its endangered status and highlighting some characteristic features. The final part includes a brief summary and conclusion.

2 Background

One characteristic of the Chinese community of Timor-Leste is its "Hakka-dominant homogeneity" (Berlie 2015: 38). I will therefore briefly introduce the Hakka people and their language. Afterwards, I provide information on East Timor, its history as well as the linguistic scene of the country. Finally, I focus on the historical relations between East Timor and Macao as Portuguese possessions during the colonial period.

2.1 Hakka

The term Hakka refers to a Sinitic language of southern China as well as its speakers, whose number is estimated to be somewhere around 40 million (Lau 2017; Chappell 2015: 15). The Hakka heartland is located at the intersection

---

2 See e.g. Pan (1999), Rae & Witzel (2008), and Tan (2013). Exceptions are Purcell (1966: 3), which briefly mentions Portuguese Timor in a table summarizing the number of ethnic Chinese in various countries of Southeast Asia; and Wurm et al.'s (1987) *Language Atlas of China*, which shows a Hakka-speaking Chinese community in East Timor on Map B 16b of "Overseas Chinese" communities worldwide. Neither work provides further information.

of the Guangdong, Guangxi, Fujian, and Jiangxi provinces (Skinner 1967: 102; Leong 1997: 23–25), in areas made up of remote, resource-poor highlands. Historically, these tended to be frontier regions which other Chinese avoided, and the Hakka's socio-economic status was low compared to other groups living in the same regions. This resulted in a high degree of geographic mobility: the Hakka are "sojourners and migrants par excellence, rarely hesitating to pull up stakes and move on to another destination, whether within or outside China" (Pan 1998: 26) and "most willing among all the Chinese migrant groups to go into the rural areas" (Wang 2003: 58). Indeed, according to Wang (2003: 222), "where the migrants need to be adventurous and act as pioneers entering undeveloped areas, the Hakka had an edge over everybody else". Outside of mainland China, Hakka speakers are widespread particularly in Taiwan and parts of Southeast Asia (cf. Wurm et al. 1987, Map B 16b). Small communities are also found across the globe, e.g. in Mauritius (Lefort 2018) and in Suriname (Fat 2015). Wang (2003: 219) notes that even overseas, the Hakka tend to "not [be] the majority where there are large concentrations of Chinese people", but "a minority among Chinese minorities", although they are the majority in some countries with relatively small Chinese populations, such as Timor-Leste or Mauritius. Even in large countries where they are a minority, the Hakka may regionally be a majority. For instance, they are estimated to make up less than 20% of Chinese-Indonesians (Lim & Mead 2011: 19). On the Indonesian island of Belitung, however, they accounted for over 70% of the total population in 1930 (Heidhues 1996: 176), and they presumably make up a large proportion of the island's population to the present day. Possibly because of their mobility, the Hakka are reported to be "unmatched in the extent of their cultural deviance, […] stubborn resistance to assimilation, and ethnic self-consciousness" (Leong 1997: 19).

According to orthodox Hakka historiography, the Hakka migrated south from the Central Plain beginning about 300 CE and arrived at their current locations in a series of five migration waves (Luo 1933). This account is intended to cement their status as "true Han Chinese from the cradle of Chinese civilization", but is "ethnic rhetoric" rather than pure scholarship (Leong 1997: 29). While a southward migration from the Yellow River Basin towards the modern provinces of Guangdong, Guangxi, Fujian, and Jiangxi did take place (Norman 1988: 181–85), this population movement preceded the split of what Norman (1988) calls Old Southern Chinese into Min, Hakka and Yue (Cantonese). The differentiation of Hakka from those groups "can presumably be dated at least from the Song" (Leong 1997: 33), i.e. between the tenth and the thirteenth centuries CE. Similar dates can be assumed for the settlement of the Hakka heartland. The name Hakka, as well as a sense of Hakka identity, however,

are much younger. The term Hakka derives from Cantonese *haak³gaa³* (客家) 'guest families', and Constable (1996) and Leong (1997) suggest that a Hakka identity was constructed after 1800, in the wake of territorial conflicts with Cantonese-speaking residents of Guangdong (cf. Sagart & Chappell 2011). Prior to this period, a variety of localized and context-specific designations were used to refer to Hakka-speaking groups by both outsiders as well as Hakka speakers themselves (cf. Leong 1997: 47, 65–67). Lau (2017) reports that as late as the 1980s there were Hakka speakers in China who were either ignorant of the term Hakka or denied their membership to this group, showing that the construction of the Hakka identity was a lengthy process.

Linguistically, various local dialects of Hakka can be distinguished. Wurm et al. (1987) counts eight subdialects, whereas Lau (2017) identifies four. There are conflicting reports regarding the extent of the differences between them: Norman (2003: 72), for instance, argues that in Hakka as well as other Sinitic languages differences between individual dialects are "in many cases considerable" (cf. Lau 2017). According to Sagart & Chappell (2011), on the other hand, Hakka dialects are remarkably homogeneous (cf. Wurm et al. 1987). What this apparent contradiction seems to boil down to is what individual scholars label as Hakka dialects, a question which is far from trivial: Sagart (1998) points out numerous problems in delimiting Hakka from Gan, and concludes that "[m]ost accounts of the geographical extent of the Hakka dialect rely, implicitly, on Hakka self-awareness, rather than on some linguistic feature" (Sagart 1998: 297). The Meixian (梅縣; or Moyen, Moi-yan) dialect of Hakka is considered the most representative of Hakka dialects and is generally taken as a *de facto* standard. Meixian city, in the northeast of Guangdong, is considered the cultural capital of the Hakka (Sagart & Chappell 2011), and according to Lau (2017), the dialect spoken there can be understood by most Hakka speakers.

2.2 *East Timor*

Timor-Leste (officially the Democratic Republic of Timor-Leste; also known as East Timor) is a sovereign state which occupies the eastern half of the island of Timor as well as the Oecussi enclave in the western part of the island. Fig. 2.1 shows its location within the Indonesian archipelago. With a surface of just over 15,000 km² (Government of Timor-Leste) and a population of 1.3 million (UNESCO Bangkok: Asia and Pacific Regional Bureau for Education 2018), it is one of the youngest nations in the world, having gained independence from Indonesia in 2002.

Timor-Leste's capital, Dili, is the country's main urban centre, with a population of 300,000; this is followed by Baucau with approximately 125,000

FIGURE 2.1   Timor-Leste in relation to its closest neighbours

inhabitants (General Directorate of Statistics). Fig. 2.2 shows the provinces and towns of Timor-Leste. Over 69% of Timor-Leste's population lives in rural areas (UNESCO Bangkok: Asia and Pacific Regional Bureau for Education 2018).

Following Hajek (2000: 213), East Timor's history since Western contact can be divided into three discrete periods:

(1)   Portuguese contact and colonization (1500s–1975)
(2)   Indonesian occupation (1975–99)
(3)   Post-Referendum/Independence (1999/2002–present)

The Portuguese arrived in the Lesser Sunda Islands in the early sixteenth century, establishing themselves on the islands of Solor, Ende, Timor, and Flores (Subrahmanyam 2012: 219; Figueiredo 2004: 113). Attracted to Timor by its rich supply of sandalwood, they initially traded the precious commodity via their base on Solor (see Fig. 2.2 below). Towards the end of the sixteenth century, they started establishing permanent settlements on Timor's north coast, and the island was officially declared a Portuguese colony in 1702 (Sousa 2006: 15).

Around the turn of the sixteenth and seventeenth century, the Dutch East India Company (VOC) began to challenge Portuguese dominance in the area, capturing important Portuguese bases like Ambon and Melaka (Hägerdal 2012: 34–35). The Portuguese-Dutch rivalry arrived in the Timor area in the

FIGURE 2.2   Map of Timor-Leste
CARTOGIS SCHOLARLY INFORMATION SERVICES, THE AUSTRALIAN NATIONAL
UNIVERSITY, LICENSE BY-SA LICENSE; DOWNLOADABLE FROM HTTP://
ASIAPACIFIC.ANU.EDU.AU/MAPSONLINE/BASE-MAPS/TIMOR-TIMOR-LESTE

mid-seventeenth century and dominated the colonial history of Timor up until the nineteenth century. It was finally concluded in 1859, when the Dutch and the Portuguese agreed on a border between their respective territories on Timor in the Treaty of Lisbon, and Portugal formally ceded claims to Solor and other nearby islands to the Netherlands (Treaty of Lisbon 1861; cf. Gunn 1999: 75). The land border between the modern states of Indonesia and Timor-Leste largely conforms to the division established in this treaty.

East Timor remained a Portuguese colony until 1975, with the dubious distinction of being "the most backward of the Western colonies in Asia" (Kammen & Chen 2019: 7). Portuguese control over the mountainous interior of the island was tenuous up until the early twentieth century and was consolidated only after several rebellions and subsequent pacification campaigns (Gunn 1999: 82–101). The colony was also used as a "dumping ground for subversives" (Moutinho 2004: 80), including both political exiles and ordinary criminals (Gunn 1999: 114). More detailed information on the colonial history of East Timor can be found in Hägerdal (2012), Subrahmanyam (2012), and Gunn (1999).

In 1975, East Timor declared independence from Portugal and was promptly invaded by Indonesia, which had stoked and exploited fears of communist tendencies in the leading Fretilin (*Frente Revolucionaria do Timor Leste*

Independente) party (see Gunn 1999: 147–54; Thaler 2012: 211). East Timorese forces fought the invasion with some success in the first few years but were ultimately overwhelmed by Indonesia's military superiority. Nonetheless, small guerrilla forces remained active throughout the 24-year occupation, which, according to Durand (2011), was "one of the greatest human tragedies of the second half of the twentieth century". It is estimated that 20%–30% of East Timor's population had died by 1999 (Durand 2011; cf. Kiernan 2003). In 1999, Indonesia allowed the East Timorese to hold a referendum in order to decide between autonomy within Indonesia or independence. When a clear majority (78.5%) voted in favour of independence (KPP HAM 2006: 37), pro-Indonesian militia retaliated, causing more than a thousand deaths, the displacement of hundreds of thousands and the near-total destruction of the country's infrastructure (Babo Soares 2003: 70; KPP HAM 2006: 54). In 2002, independence was finally achieved under the direction of the United Nations, who remained in the country until 2012 (Martin & Mayer-Rieckh 2005; United Nations 2019). The situation in Timor-Leste has since been fairly stable, apart from a period of unrest from 2006 to 2007 (Scambary 2009). The country is, however, still very much dependent on foreign investment and aid (The Heritage Foundation 2019).

Timor-Leste has two co-official languages, the indigenous lingua franca Tetun Dili and Portuguese. The country is characterized by considerable linguistic diversity, with some 20 Austronesian and Papuan (Timor-Alor-Pantar) languages being spoken according to Glottolog (Hammarström et al. 2019).[3] Neither a number nor individual language names apart from Tetun and Portuguese are given in Timor-Leste's Constitution,[4] which merely states that "Tetu[n] and the other national languages shall be valued and developed by the State". In the linguistic literature on Timor-Leste, Hakka generally goes unmentioned; exceptions are, to my knowledge, only Hull (2002) and Hajek (2000). Hajek (2000) is a brief language history of East Timor which includes

---

3 Different counts are found in other sources: for instance, up until approximately 2010 Timor-Leste's *Instituto Nacional de Linguística* gave the number of indigenous languages as 16 (cf. Hull 2002). The most recent edition of Ethnologue (Eberhard, Simons & Fennig 2019) lists 21 languages, and Edwards & UBB's (2018) linguistic map of Timor shows 26 languages spoken in Timor-Leste. The differences are mostly due to different decisions on what is to be counted as a separate language and what as a dialect of another language. Ethnologue and Glottolog furthermore include an extinct Portuguese-based creole (Bidau Creole Portuguese), and Ethnologue adds Portuguese, which, despite its status as an official language, is not commonly used in everyday life.

4 Downloadable from http://timor-leste.gov.tl/?cat=37&lang=en.

a paragraph on the Chinese minority (2000: 218), but does not give any linguistic information. Hull (2002), formerly the director of the *Instituto Nacional de Linguística*, gives very basic information on the 16 indigenous languages of Timor-Leste and, presumably in order to illustrate relatedness, provides the numerals from one to ten in each language. He also includes a paragraph on the non-indigenous languages of the island, where two lines are devoted to Hakka. Both sources appeared at a time when Timor-Leste had gained worldwide attention thanks to the 1999 referendum, when little was known about its languages. Much progress has since been made in the description of Timor's indigenous languages. Our knowledge of Timor-Hakka, however, remains as rudimentary as it ever was.

2.3   *The Portuguese in Asia: Timor and Macao*

As Portugal's influence in Asia diminished, increasingly close ties developed between the Portuguese territories in the Lesser Sunda Islands – Timor, Solor, and a few nearby islands – and Macao, Portugal's other Far Eastern possession. The Portuguese had established a trading post, which rapidly developed into an important centre of commerce, in Macao in 1557. Of particular importance was the silk and silver trade with Japan, in which the Portuguese acted as middlemen between the Chinese and the Japanese (Flynn & Giraldez 1996: 56). There were also important connections to Southeast Asian ports, in particular Melaka and Manila (Ptak 2006). In the seventeenth century, a series of events took place that dramatically affected these connections: first, the Portuguese were expelled from Japan in 1639, due to the latter's concerns about Portuguese missionary activities on the islands (Subrahmanyam 2012: 179). Then, in 1640, the Portuguese Restoration War ended the dynastic union between Spain and Portugal (Disney 2009). This effectively ended trade between Macao and Manila, which were in Portuguese and Spanish hands, respectively. Finally, in 1641, the Dutch seized Melaka from the Portuguese (Subrahmanyam 2012: 183). Thus, within three years, Macao was deprived of its three most important markets. By the 1670s, "the Macau-based traders [were] a pale shadow of the prosperous mercantile community of the early years of the same century", "searching constantly for outlets which the Dutch might permit them" (Subrahmanyam 2012: 222).[5] As a result, links with Portuguese territories such as Timor, which had previously been of minor importance, became more

---

5   See Subrahmanyam (2012) for a more detailed analysis of the decline of the *Estado da Índia* in the seventeenth century.

prominent. In fact, by the end of the seventeenth century, Timor, Solor, and some other islands in the region were the only remaining Portuguese possessions in the Far East apart from Macao. The links between the Portuguese-ruled Lesser Sunda Islands and Macao were administrative, commercial and economical, and religious. Following Sousa (2006), these are briefly discussed in turn.

Until 1844, Macao and Timor and Solor were separate territories under the jurisdiction of the Portuguese *Estado da Índia*. Even so, the government of the captaincy of Timor and Solor was heavily dependent on the authorities in Macao by 1836. For instance, the two territories were represented in the Parliament of Lisbon by a single ambassador (Sousa 2006: 17), and it is reasonable to assume that this person generally came from Macao. In 1844, the two Far Eastern territories were joined to form a new province, independent from the *Estado da Índia*. The seat of government for the *Provincia de Macau, Solor e Timor* was in Macao; a subordinate governor resided in Dili. In 1850, Timor and Solor were declared an autonomous province. This was a period of intense territorial conflict with the Dutch, and it is likely that the elevation of the islands to the status of a separate province was intended to give the governor the power to conduct the negotiations with the Dutch that were to lead to the 1859 Treaty of Lisbon (Sousa 2006: 17). In 1851, Timor and Solor went back under the jurisdiction of Macao, and in 1856, the Far Eastern territories were subordinated again to the *Estado da Índia*. Timor again briefly became a separate colony in 1863 before being made an autonomous district within the newly formed *Província de Macau e Timor* in 1866. In 1897, Timor became independent from Macao, and from then on, it was directly administered by Lisbon. Table 2.1 sums up the administrative status of Portuguese Timor from the time it officially became a Portuguese colony in 1702 until its declaration of independence in 1975 (excluding the period between 1942–45, when it was occupied by Japan).

Regular trade relations between Macao and Timor existed since the late 1500s (Pinto 2014: 151). In the seventeenth century, Macao imported sandalwood, slaves, turtle shells, honey, and beeswax from Timor, while Timor imported textiles, metal, gold, as well as rice purchased in Batavia from Macao traders (Sousa 2006: 18). As Macao's more lucrative markets were lost around 1640, the relative importance of Timor for the Portuguese trading network increased. In the late seventeenth century, the viceroy of Portuguese India granted Macao the monopoly on sandalwood trade from Timor, and the Timorese were banned from selling their main export commodity to the Dutch. During this period, two to three boats went from Macao to Timor annually, stopping at a

TABLE 2.1   Administration of Portuguese Timor (1702–1975)

1702–1844	Captaincy of Timor and Solor under the jurisdiction of the *Estado da Índia*
1844–1850	Province of Macao, Solor and Timor
1850–1851	Province of Solor and Timor
1851–1856	Under the jurisdiction of Macao
1856–1863	Under the jurisdiction of the *Estado da Índia* (jointly with Macao)
1863–1866	Province of Timor
1866–1896	Province of Macao and Timor
1896–1975	Variable status as province or autonomous district under the jurisdiction of Lisbon

number of Dutch ports[6] in the Indonesian archipelago along the way (Sousa 2006: 18). Next to the official channels of trade, there were also varying degrees of trade between Timor and Macao organized by private merchants from Macao (Pinto 2014). Even so, the Timor market was never a very lucrative one, and the volume of trade with Macao was relatively low. Towards the end of the eighteenth century, Timor's sandalwood reserves were depleted. The colony kept running deficits, and Macao supported it with yearly subsidies from 1811 onwards (Gunn 1999: 52), even in periods when there was no direct administrative link between the two territories. Records from the period when Timor and Macao were joined in the *Província de Macau e Timor* in 1866 show that these subsidies were unpopular with the population in Macao. In the same period, Macao military personnel were obliged to two years of service in Timor (Sousa 2006: 17).

Macao and Timor were also tightly linked by religious connections, in particular during the 1800s. In the early part of the century, missionaries were sent from Macao to Timor to help promote Catholic religious education in Timor. In 1864, Timor was made part of the diocese of Macao and remained part of that diocese until 1940. Missionary efforts were renewed towards the end of the nineteenth century, when missionaries, priests, and nuns were sent from Macao to Timor (Sousa 2006: 20–21). While many of them were Portuguese, there were also some Chinese people among them. For instance, Sousa (2006: 21) reports that a group of missionaries and priests arriving from Macao in

---

6   E.g. Batavia, Melaka, Madura, Bali, and Larantuka.

1877 included a Cantonese priest, whose focus was on converting the Chinese of Timor to Catholicism.

The various administrative, commercial, and religious links may furthermore have resulted in a linguistic connection between Timor and Macao: Bidau Creole Portuguese, spoken in Dili until the twentieth century, exhibits a number of striking similarities with Macao Creole Portuguese and was likely influenced by it (Baxter & Cardoso 2017: 297).

3   The Chinese-Timorese Community

Throughout most of the nineteenth century, both the Portuguese and the Dutch had limited control over the Timorese hinterland. In this period, the island formed an "integrated whole" with respect to Chinese business activities (Kammen & Chen 2019: 21). By the early twentieth century at the latest, the ethnic Chinese of West Timor and East Timor can be considered separate communities. The extent to which they have developed along separate ways historically and linguistically is a topic for further research. This section highlights, successively, the history of the Chinese community of East Timor from the colonial period, the Indonesian occupation, to the present day. More in-depth accounts can be found in Kammen & Chen (2019) and Soares (2019).

3.1   *Origins and the Colonial Period*
In the fourteenth century, Timor was the southernmost endpoint of Chinese trading voyages in Southeast Asia (Gunn 2016: 128; Kwartanada 2001: 2), and Timor is mentioned as a place rich in sandalwood in a thirteenth century Chinese source, Zhao Rugua's 趙汝适 account of countries outside China and the products they traded (Hirth & Rockhill 1911: 208–9). According to Ptak (1987; quoted in Sousa 2006: 14), trade relations between China and Timor may date back as early as the twelfth century. In any case, by the time the Portuguese arrived in the region, Chinese traders were "an accepted part of the scene in the waters round Timor" (Ormeling 1956).

During the colonial period, the most prominent Chinese settlements on Timor were in Kupang, Lifau, and Dili. Near the island's western tip, Kupang was the site of a Portuguese settlement established in 1646 (Pinto 2014: 144), but was captured by the Dutch only 7 years later (Jannisa 2019: 48). Today, it is the capital of Indonesia's Nusa Tenggara Timur province. Lifau, located in Timor-Leste's Oecussi enclave, was the seat of government of Portuguese Timor from 1702 until 1769 (Jannisa 2019: 47). At that point, the colony's capital was moved to Dili. The earliest reports of a Chinese community living on Timor

come from Lifau: in 1699, a British naval officer who passed by the island made note of a few Chinese living there, and a few years later, in 1702, the ordinance of the Portuguese judicial officers of Lifau distinguished Portuguese, Timorese, and Chinese residents for tax purposes (Pinto 2014: 158–59). In the early eighteenth century, Chinese traders also began to settle in Kupang (Hägerdal 2010: 23). By 1770, Kupang had a Chinese neighbourhood complete with a Chinese temple (Lombard-Jourdan & Salmon 1998: 396). It is unclear when the first Chinese settled in Dili, although this city appears to have risen to prominence as a centre of Chinese activity only after it was made the capital of Portuguese Timor. Records show that in 1785 the colony's governor publicly invited Chinese traders to settle in the new capital in an effort to stimulate its economy. A few years later, he reported that seven Chinese traders had established themselves in Dili, and a *capitão Cina*[7] existed by the early nineteenth century (Pinto 2014: 159; cf. Kammen & Chen 2019: 52). According to Figueiredo (2004: 228), 20 Chinese lived in Dili's Bidau neighbourhood in 1812.

Immigration from China to Portuguese Timor followed two patterns, an official and an informal one (Pinto 2014). The official channel linked East Timor to Macao. Orchestrated by the colonial authorities, it mainly concerned "the trade in sandalwood and the Macao government's occasional efforts to bring craftsmen and workers to the island" (Pinto 2014: 161). Efforts to boost the remote colony's economy by encouraging the immigration of Chinese merchants and craftsmen started in the late eighteenth century and continued throughout the nineteenth century. Portuguese Timor also served as a place of exile for convicts from Macao. It was among "the most challenging tropical locations" (Coates 2018: 49–50) in the Portuguese empire,[8] and to be exiled to Timor was reserved as punishment for the most serious crimes. In 1844, a report in the Macao newspaper *A Aurora Macaense* criticized the long-standing practice of "dumping its most incorrigible, uncouth, and uneducated convicts" (Gunn 1999: 52; cf. Sousa 2006: 21) in Timor. This form of punishment was practiced at least from the first half of the nineteenth century until after the Second World War (Kammen & Chen 2019: 22, 110).

The informal channel involved private individuals, departing from a variety of ports (including Macao, Swatow, and Hong Kong) and travelling to Portuguese Timor by different routes (Kammen & Chen 2019: 54–56). This

---

7   This title referred to an official tasked with overseeing the community and liaising with the colonial government. This system was copied from the Netherlands Indies (Kammen & Chen 2019: 52) and can be traced back to pre-colonial times (Ooi 2004).

8   Recall also that Kammen & Chen (2019) described Timor as "the most backward of European colonies" (cf. section 2.2).

channel became significant around the middle of the nineteenth century, when a series of upheavals in southern China (the Opium Wars of the 1840s, severe floods between 1846–48, the Taiping Rebellion and the Punti-Hakka Clan Wars in the 1850s and 1860s) caused emigration from the region to soar. While Portuguese Timor was by no means a prominent destination, a portion of the migrants found their way there (Kammen & Chen 2019: 35–36), and the informal channel soon overtook the official one, contributing a much larger number of Chinese immigrants.

Throughout the eighteenth and nineteenth centuries, the volume of migrants brought to Portuguese Timor via both channels remained on a low level. Sousa (2006: 18) reports that by around 1800, some 300 Chinese families, mainly from Macao, were living scattered on the island; in other words, this number included both the Portuguese and the Dutch parts of Timor. The Chinese community of Portuguese Timor started to increase dramatically in the early twentieth century (Kwartanada 2001: 4; Kammen & Chen 2019: 56–57). This surge was caused by two events, namely the overthrow of the monarchy in Portugal in 1910, and the end of the Qing Dynasty in China in 1911. In China, the turmoil following the overthrow of the Qing resulted in a wave of emigration to Southeast Asia. The establishment of the Republic in Portugal resulted in administrative and economic reforms of the empire which facilitated Chinese immigration (Kammen & Chen 2019: 81–82). The Chinese population of Timor continued to increase in the post-WWII period, both through high birth rates as well as immigration (Kammen & Chen 2019: 110, 114).

Fig. 2.3 illustrates the growth of the Chinese community of Portuguese Timor from the late nineteenth century up to 1970; the numbers on which this graph is based are taken from Kammen & Chen (2019: 29, 38, 70). The figures for 1907, 1927, 1950 and 1970 come from colonial censuses, while the nineteenth century figures are estimates. Despite the considerable growth in the twentieth century, the community remained small both in absolute numbers as well as relative to the colony's total population: in 1927, ethnic Chinese residents made up 0.37% of Portuguese Timor's population; in 1950, 0.71%; and in 1970, approximately 1.0%.[9] It must be noted, however, that the figures for the Chinese population of the colony may be somewhat too low, as some Chinese may have tried to "dodge the census taker so as to avoid payment of the annual fee for foreigners" (Kammen & Chen 2019: 71).

---

9   The total population of Portuguese Timor was 451,604 in 1927 (Telkamp 1975: 6–7 quoted in Kwartanada 2001: 5), 442,378 in 1950, and 609,477 in 1970 (Telkamp 1979: 75). Again, the figures come from colonial censuses.
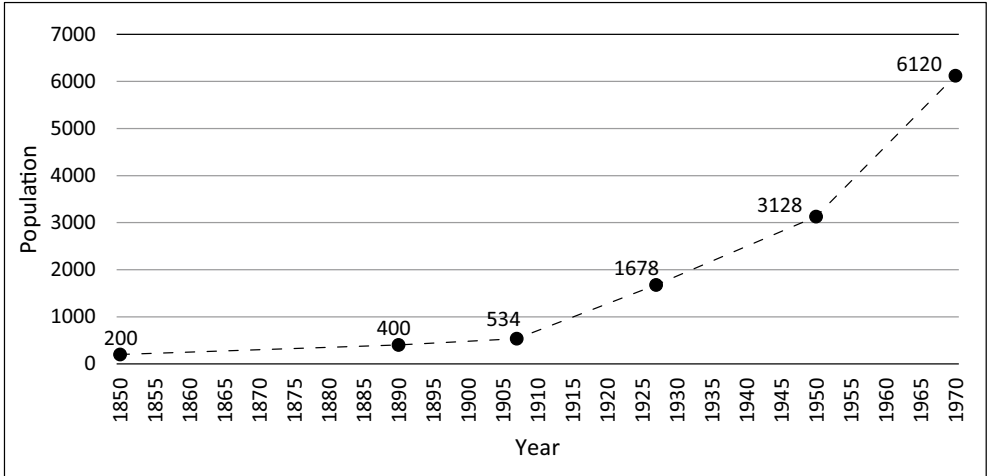
FIGURE 2.3   The East Timorese Chinese community in numbers, 1850–1970
            OWN WORK

Despite their small number, the ethnic Chinese dominated the economy of Portuguese Timor. Several early sources stress that the Chinese-Timorese "were the most useful part of the population" (Gunn 1999: 60), or "the only part of the population which carries on trade" (Kwartanada 2001: 3; cf. also Sousa 2006: 19; Pinto 2014: 157). By the mid-1800s, the Chinese in Dili controlled the small and medium retail business as well as a significant part of the export business (Pinto 2014: 163), and in the early 1960s, all but three or four of some 400 plus wholesale and retail businesses in the colony were owned by ethnic Chinese businessmen (Dunn 1983: 9). The Chinese-Timorese were also "the main brokers of the grain and coffee markets" (Kwartanada 2001: 7). Beyond trade and commerce, Chinese craftsmen such as stonemasons and carpenters were highly sought after by colonial officials, who bemoaned the shortage of skilled labour among the indigenous East Timorese (Sousa 2006: 19). It appears that many of the Chinese craftsmen were contract labourers, a large proportion of whom presumably returned to China after the end of their contracts (Kammen & Chen 2019: 35, 54).

Throughout the colonial period, ethnic Chinese businessmen served as intermediaries between the indigenous population and the Portuguese as well as other parts of the outside world. Kwartanada (2001) characterizes them as a middleman minority, i.e., a minority group "serving an intermediary position between the majority group and other segregated minority groups" (Douglas & Sáenz 2007: 147). Middleman minorities are usually sojourners who migrate

for economic reasons and "tend to concentrate in certain occupations, notably trade and commerce" (Bonacich 1973: 583). Many Chinese merchants sought to strengthen good business relationships with Timorese chiefs by marrying women from their families (Nicol 2002: 31). The Chinese were also closely associated with the colonial powers (see Pinto 2014: 159; Nicol 2002: 31), who relied on them not only for their personal needs, but also for the collection of head taxes from the indigenous population in the rural areas of East Timor. As such, they were "a condition for [the state's] very existence and ability to function" (Kammen & Chen 2019: 98).

Despite their reliance on them, the Portuguese colonizers' attitude towards the Chinese-Timorese minority was marked by distrust. The Portuguese aimed to restrict and control Chinese immigration and trade, and Chinese-Timorese merchants were often at risk of falling "victim to capricious colonial officials and military personnel" (Kammen & Chen 2019: 42). Discriminatory practices continued into the twentieth century and included not being allowed to loiter in the streets after 8:00 pm; needing a permission for slaughtering animals and a license to employ workers; and restrictions on travel and visa requirements (Kammen & Chen 2019: 96, 123). In response, the Chinese community developed a high degree of solidarity and a strong sense of identity (Kammen & Chen 2019: 82, 98; Wise 2006: 154). A system of social institutions such as business and self-help associations "entirely out of proportion to [the community's] size" offered "protection from the most repressive expressions of state authority" (Kammen & Chen 2019: 98). There was also a well-established school system, with at least one Chinese school in every district by 1933 (Kammen & Chen 2019: 88). Up to 1975, Chinese-Timorese children generally attended Portuguese schools as well as Chinese schools where they were available. A Chinese-Timorese woman quoted in England (1999) reports that "[b]efore 1975, I went [to a Chinese school] in the mornings and to the Portuguese school in the afternoons". The Chinese schools in Portuguese Timor were supported with textbooks and teachers from Taiwan (Chew & Huang 2014: 313) and China (Wise 2006: 147), and instruction took place either in Hakka or in Mandarin (cf. Chew & Huang 2015: 313).

Among the Chinese cultural institutions in Dili were also a temple and a cemetery. A first temple was destroyed around 1915. A new one, the Guandi temple 關帝廟, was built in 1928 and still exists today (Timor Tourism 2015). A part of the Chinese-Timorese community practiced Chinese folk religion and Buddhism, while, as mentioned previously, substantial numbers also adopted Catholicism. In many instances, the belief systems existed side-by-side. For instance, some families attended both the Chinese temple as well as Catholic church services (Raynor 2019; cf. Chew & Huang 2014: 317). In

the Chinese cemetery in Dili, both graves with Christian symbols as well as those invoking deities from Chinese folk religion can be found. Berlie (2015: 42) notes in particular the importance of Tushen 土神, the genius loci of Chinese popular religion, in the cemetery, which he takes as evidence of how the Chinese-Timorese have taken root in Timor (cf. Tan 2018). In general, a Chinese-Timorese culture developed which was characterized by "mixed cultural forms influenced by a long period of Portuguese colonial rule on East Timor" (Chew & Huang 2014: 301–2). Notably, however, "forms of social organization characteristic of overseas Chinese elsewhere in Southeast Asia were far less important in Timor" (Kammen & Chen 2019: 85); for instance, surname associations were practically absent.

An early nineteenth-century source notes that the Chinese-Timorese community of Dili was of mixed origin (Pinto 2014: 159). According to Kammen & Chen (2019: 37–8, 54–5, 59), three main linguistic groups of Chinese immigrants can be identified: Hokkien, Hakka, and Cantonese. The linguistic background of these groups tends to correlate with some social characteristics. Hokkien speakers ("Baba"; see Kammen & Chen 2019: 37) made up the majority of the earliest immigrants, who arrived in Timor in the eighteenth and early nineteenth century, well before the colonial spheres of dominance were fully consolidated. They were centred in Kupang, in today's West Timor. From there, they spread along to the north coast of the island to Dili and other towns in Portuguese Timor. Most were merchants and presumably made their way to Timor via the informal channel of immigration.

Cantonese speakers from the Pearl River Delta arrived in relatively small numbers either as skilled labourers or as convicts. As noted above, efforts to bring merchants and craftsmen to Portuguese Timor started in the late eighteenth century. Convicts were sent to Portuguese Timor from at least the first half of the nineteenth century up until the 1950s. In the twentieth century, they made up the majority of Cantonese speakers arriving in the colony (Kammen & Chen 2019: 66–67). Skilled labourers were mostly on short-term contracts and left Timor when their contracts ended. Although they were at the bottom of the social ladder within the Chinese-Timorese community, convicts often stayed on, finding jobs as carpenters or construction workers (Kammen & Chen 2019: 134).

Hakka speakers started arriving in the second half of the nineteenth century, many fleeing from upheavals in China. By the twentieth century, the majority of Chinese immigrants to Portuguese Timor were Hakka speakers. There were two distinct Hakka groups, those from Xiangshan County 香山縣 (near Macao), and those from Jiayingzhou 嘉應州 (Meixian) in north-eastern Guangdong, who brought with them distinct Hakka dialects. The two groups

TABLE 2.2   Estimated percentage of Chinese in Portuguese Timor by speech-group

| Year | Hokkien | Hakka | | Cantonese |
		Xiangshan	Meixian	
1900	20–25%	25%	25–30%	25%
1915	15%	20%	40%	25%
1940	10%	15%	55%	20%
1975	–	10%	75–80%	10–15%

FROM KAMMEN & CHEN (2019: 79)

were present in Timor in roughly equal numbers at the turn of the twentieth century, but by the end of the colonial period, Meixian Hakka made up the vast majority of the Chinese-Timorese community overall. Table 2.2 illustrates how the make-up of the community changed over time.

Up to the early twentieth century, the Chinese-Timorese community was mostly male. Many of the early Hokkien-speaking immigrants thus married Timorese women. Marriages with women from local ruling families were also an important means of establishing business relations (Kammen & Chen 2019: 43). Situated at the bottom of the social ladder, Cantonese-speaking convicts were likely to remain single or marry Timorese women (Kammen & Chen 2019: 134, 145). Only in the twentieth century, a sufficient number of Chinese women arrived to allow the community to consolidate (Kammen & Chen 2019: 79). Most were Hakka and tended to marry other Hakka speakers. Mixed marriages were also more likely to happen in rural settings: Chinese men who established themselves in rural areas were more likely to marry East Timorese women and became more integrated in East Timorese society and culture (Wise 2006: 150). It was mostly Hakka speakers who ventured beyond the urban centres, while the majority of Hokkien and Cantonese speakers remained in the cities (Kammen & Chen 2019: 73–74).

3.2 *The End of the Colonial Period and the Indonesian Occupation*

By the end of the colonial period, Dili had become the centre of Chinese cultural and institutional life in East Timor. However, several major district seats also had significant numbers of Chinese residents (Kammen & Chen 2019: 82, 74), and there were few rural areas which "did not at least have one Chinese shopkeeper" (Hajek 2000: 218). The inhabitants of the remote subdistrict seat of Iliomar on the south-eastern coast of Timor-Leste, for instance, still

remember the Chinese shopkeeper known as *Sina Fernando*[10] who used to live there; his shop is now in ruins. As shown in the introduction, the Chinese were very much an accepted part of East Timorese life and had found their way into ritually significant expressions of indigenous East Timorese culture. There are no reports of open animosity between the East Timorese and the Chinese-Timorese, and it appears that "Chinese cultural issues and identity posed no problems to the other groups settled in and around Timor" (Pinto 2014: 158). Nonetheless, the relations between the two groups were not without problems. Some East Timorese resented the self-imposed social isolation of the Chinese-Timorese community (Nicol 2002: 58), and jealousy over their economic success can also be assumed to have been widespread (cf. Kammen & Chen 2019: 157). Chinese-Timorese were furthermore accused of a "lack of commitment" to East Timor. For instance, many Chinese-Timorese opted for ROC citizenship rather than Portuguese citizenship. According to stories circulating among East Timorese expats in Australia, it was only when they hoped to gain the status of East Timorese refugees that the Chinese-Timorese referred to themselves as East Timorese and tried to obtain Portuguese passports (Wise 2006: 157; Kammen & Chen 2019: 134–35).[11]

During the colonial period, the Chinese-Timorese community had steadily increased in size. The events surrounding the Indonesian occupation of East Timor resulted in a dramatic reduction of the community. Assuming a constant growth rate and taking into account the fact that the actual population was likely somewhat larger than what is reported in official figures mentioned previously, Kammen & Chen (2019: 72) estimate that some 7,500–8,000 ethnic Chinese people lived in Portuguese Timor by 1975. It must be noted, however, that other sources give – sometimes drastically – different estimates. For instance, Dunn (1983: 9) quotes a lower number of 6,170, although he concedes that this figure omits the so-called Chinese-Timorese mestizos "who considered themselves Chinese in cultural terms". A number almost twice that given in Kammen & Chen (2019), 14,000, appears in Telkamp (1979: 76). An even higher estimate is given by Berlie (2015: 40), according to whom the community numbered up to 25,000 people, and Miao et al. (2015) quote a number as high as 30,000. The total population of Portuguese Timor at the time is estimated to have been about 700,000 (Kwartanada 2001: 5; Telkamp 1979: 75).

---

10   *Sina* refers to both China as well as Chinese people.

11   Hostility directed at middleman minorities from the host society is common (Bonacich 1973: 589–93). In particular, accusations of disloyalty to the host country, or dual loyalty, are a recurrent theme (Bonacich 1973: 591).

At 8,000, the Chinese-Timorese would thus have accounted for 1.5% of the colony's population.[12] Although this is a clear increase from 1970, when they made up 1% as mentioned above, the Chinese-Timorese still remained a tiny minority in East Timor.

The decline of the Chinese-Timorese community set in even before the Indonesian invasion, namely in the time leading up to Portuguese Timor's declaration of independence in 1975. In this period, *Fretilin*, as the strongest of the newly formed political parties, called for "an end to foreign exploitation and a people's economy based on self-sufficiency and cooperatives" (Kammen & Chen 2019: 152), and some anti-Chinese sentiments were expressed (Nicol 2002: 60). Many Chinese-Timorese were thus worried about their prospects in an independent Timor-Leste. In addition, memories were still fresh of Indonesia's Communist Purges in 1965–66, during which many Chinese-Indonesians had lost their lives. As a result, as many as 2,000 Chinese-Timorese fled East Timor in this period (Kammen & Chen 2019: 155; Nicol 2002: 60).

During the 1975 invasion and the ensuing 24-year occupation mentioned in the previous section, Indonesian forces reportedly specifically targeted Chinese-Timorese (Kiernan 2003: 202). According to Dunn (1983: 285–86), some 700 Chinese-Timorese were killed in Dili in the first days of the invasion, and the Chinese communities of several other towns were practically wiped out. However, based on eyewitness accounts, Kammen & Chen (2019: 160, fn. 65) conclude that 700 Chinese-Timorese did not lose their lives in Dili in the invasion. It is true that the Chinese-Timorese inhabitants of the Lecidere, Colmera and Taibesi districts of Dili found themselves "near the front lines and drew the immediate attention of Indonesian military personnel", but in most cases, the Indonesian soldiers left "Chinese men, women, and children frightened but unhurt" (Kammen & Chen 2019: 159). The vast majority of victims in Dili were in fact indigenous East Timorese, not Chinese-Timorese.

Even if the death toll of the invasion was lower among the Chinese-Timorese than previously reported, life under Indonesian rule became exceedingly difficult. They were viewed with distrust by the Indonesian occupiers, who associated "the Chinese with communism, and communism with Fretilin" (Kammen & Chen 2019: 162). Due to a lack of business opportunities, many were forced to live off of their savings (Kammen & Chen 2019: 162). The Indonesian government "prohibited any activities that were culturally Chinese" (Kammen & Chen 2019: 160). In addition, all Chinese schools were closed. However, a consultant reports that clandestine Mandarin classes continued for a while and were attended by a significant number of Chinese-Timorese children (Lay

12   Or up to 4.3%, assuming Miao et al.'s (2015) estimate of 30,000.

Li Sum p.c., 14 April 2020). Religious activities were still permitted, and Dili's Guandi temple emerged from the invasion unscathed (Fong Kui Kong p.c., 16 August 2017; Miao et al. 2015).

In response to the difficult conditions in East Timor, those who had the means to do so fled the country. By 1980, an Indonesian state census showed that no more than 2,000 ethnic Chinese remained in East Timor. They were presumably stateless, that is, they had neither ROC nor Portuguese passports or lacked the financial resources to leave, or were "bound to Timor by marriage, political commitments, or fateful accidents" (Kammen & Chen 2019: 163).

A second wave of emigration took place in 1999, when, in retaliation for the vote in favour of independence, pro-Indonesian militia went on a rampage through the country. It is reported that, again, the remaining Chinese-Timorese and their residences were specifically targeted. During the 24-year Indonesian occupation, 90% of Chinese-Timorese are estimated to have left East Timor, with the majority fleeing to Australia and Macao (Wise 2006: 147; Sousa 2006: 17); others made their way to Portugal or Singapore, and some also went to West Timor or other parts of Indonesia (Kammen & Chen 2019: 163 Lay Li Sum p.c., 21 August 2017; Fong Kui Kong p.c., 16 August 2017).

3.3   *Independent Timor-Leste*

The exact number of Chinese-Timorese living in Timor-Leste today is difficult to determine, as the population censuses conducted by the country's General Directorate of Statistics do not provide information on ethnicity or ancestry.[13] According to Berlie (2015: 41), 2,400 Chinese-Timorese were registered with the Hakka association of Timor-Leste in the 2010s. A newer estimate is found in Soares (2019: 317), who offers a number of some 4,000 for 2014. Soares (2019: 315) furthermore estimates that between 4,500 and 5,000 new Chinese migrants lived in Timor-Leste at the time of writing. It is unclear whether this number refers to post-independence newcomers from China only, or also Chinese-Indonesians who migrated to East Timor in significant numbers during the Indonesian occupation (Soares 2019: 317). It does show, however, that the established Chinese-Timorese community may well be outnumbered by more recent migrants of Chinese descent.

The rather substantial increase in community size from 2,400 to 4,000 between ca. 2010 and 2014 is most likely owed to the post-independence return of Chinese-Timorese people who had left the country during the Indonesian occupation. Whether this rising trend has continued since 2014 is questionable.

---

13   However, the censuses contain information on mother tongue, which is discussed in section 4.
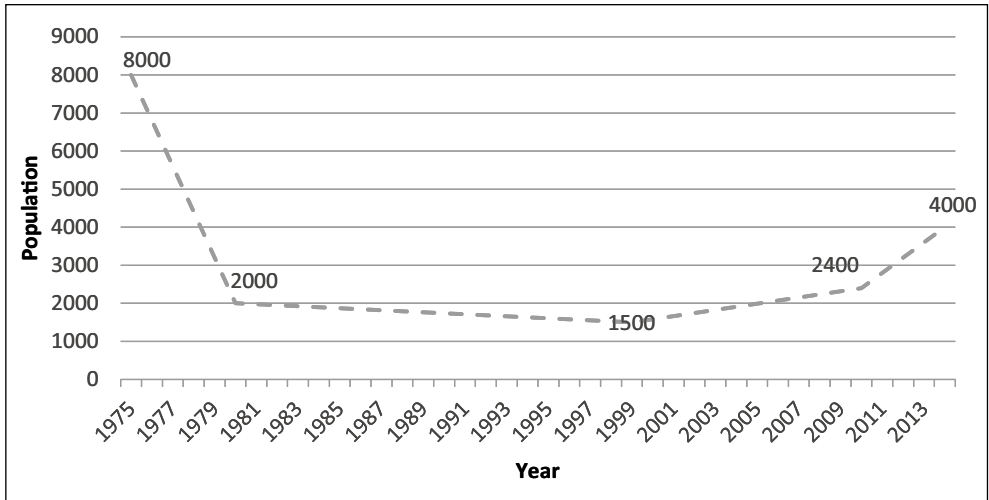
FIGURE 2.4   The East Timorese Chinese community in numbers, 1975–2014
           OWN WORK

It is likely that the majority of Chinese-Timorese who were willing to return to Timor-Leste had already done so by 2014.

Fig. 2.4 illustrates the population development since 1975. As discussed above, the Indonesian invasion had a dramatic impact on the Chinese-Timorese community.[14] In absolute terms, the post-independence growth brought it back to approximately half of the pre-invasion level. In relative terms, however, it is at a level close to that of 1927 as we saw above: according to Timor-Leste's 2015 census (General Directorate of Statistics 2015), the country's total population was 1,183,643. The 4,000 or so Chinese-Timorese thus accounted for a mere 0.34%.

The Chinese-Timorese community today is small and close-knit: as England (1999) reports, "everyone is a little bit family". The most prominent surnames are Lay, Lai and Lee (黎, 賴, 李),[15] Lim 林, Vong 黃, and Chung 鍾. To the present day, the Chinese-Timorese are dominant in the retail industry. Many are involved in construction work as government contractors or operate petrol stations and hotels. They tend to be among the wealthy in Timor-Leste, and

14   The 1975 and 1980 figures are from Kammen & Chen (2019). I have been unable to find a figure for the second wave of emigration which took place in 1999 (see section 2). The number of 1,500 in Fig. 2.4 is my own rough guess.

15   Before the standardization of romanization systems, the same name could be romanized in different ways and at the same time, a given romanization could conflate different names (Kammen & Chen 2019: xii).

it is very common for families to send their children abroad to study. There they will generally stay with relatives; indeed, everyone seems to have family abroad, especially in Australia. Most Chinese-Timorese people I spoke to had themselves spent at least some time abroad as refugees. A significant proportion of young people has been educated abroad. Many of them returned to Timor-Leste to help their parents in their businesses or to take care of them.

In contrast to Indonesia, where attitudes toward the Chinese are often marked by resentment and racism (cf. England 1999), the relations between the Chinese-Timorese and indigenous East Timorese are mostly amicable (Pinto 2014: 158; Berlie 2015: 40–43; Mulyanto 2019). In the face of the increasing numbers of new Chinese migrants, the consensus among the East Timorese appears to be that "[m]ost Chinese are all right, [...] but local Chinese are best" (England 1999). A Chinese-Timorese man quoted in England (2001) confirms that "it's easy to do business here because I am a local Chinese". The Guandi temple in Dili is "indicative of how well the ethnic Chinese and their culture have been accepted by the society" (Mulyanto 2019). The temple is visited not only by ethnic Chinese, but also by indigenous East Timorese who seek blessings or have their fortunes read (cf. Mulyanto 2019; Miao et al. 2015). During my visit in 2017, the head of the temple, a Chinese-Timorese man, was supported by two indigenous East Timorese assistants who mainly performed janitorial duties, but also helped guide and instruct visitors to the temple when it was particularly busy. Since independence, Timor-Leste has had two Chinese-Timorese ministers: Pedro Lay (sometimes Pedro Lay da Silva) was Minister of Infrastructure between 2007 and 2012 and Minister for Transport and Telecommunication between 2012 and 2015. His brother, Francisco Kalbuadi Lay, was Minister of Tourism (later Minister of Tourism, Arts and Culture) between 2012 and 2017.

Nevertheless, there are reports of animosity towards the Chinese-Timorese from the indigenous East Timorese majority. Discrimination may take many forms, from "subtle over-charging for vegetables in the market [...] and insults hurled from bored youths street-side to more sinister acts of violence" (Raynor 2019). The appearance of two recent news articles addressing this problem (Raynor 2019; Mulyanto 2019) seems to suggest that it is increasingly being acknowledged. In a Facebook comment to Mulyanto (2019), Timor-Leste's former president and co-recipient of the 1996 Nobel Peace Prize, José Ramos-Horta, praises the article and states that the Chinese-Timorese "are very Timorese in every sense". Echoing the Portuguese colonial governors of past centuries, he also welcomes new Chinese immigrants: "In Timor-Leste, they are very welcome, we do not fear them, they are energising [sic!] our country's economic life". According to Wise (2006: 158), members of the East Timorese

diaspora in Australia generally express the sentiment that it is fine for the Chinese-Timorese to practice their own culture, but their primary allegiance should clearly be to Timor-Leste.

In the early 2000s, Timorese politicians, including Ramos-Horta and former resistance leader Xanana Gusmão, visited the large Chinese-Timorese expat community in Australia in order to encourage them to return to Timor-Leste and help build the young country's economy (Wise 2006: 158). While some have returned, a large number remains overseas. Wise (2006: 147) estimates that out of the 90% of Chinese-Timorese who have left Timor-Leste during the occupation, more than two-thirds went to Australia; and according to Chew & Huang (2014: 306), the majority of Chinese-Timorese lived in this country at the time of writing.

According to Australia's Department of Immigration and Citizenship (2016), 5,706 Timor-Leste-born immigrants (48.6%) in the country report having Chinese ancestry, and 44.7% speak Hakka at home. The biggest Chinese-Timorese communities are found in Melbourne, Sydney and Darwin; smaller communities are in Brisbane and Perth (Wise 2006: 147). Outside Timor-Leste, there is a strong sense of identification as specifically Chinese-Timorese: Wise (2006: 154) reports that the Chinese-Timorese in Australia go to "extraordinary lengths" to get together, and that they prefer to "mix with other Chinese-Timorese rather than Chinese communities at large or other Timorese communities". This, she argues, is evidence of the "sense of difference from other Chinese and East Timorese and identification with East Timorese Chineseness" (cf. Hajek & Goglia 2019). Likewise, Chew & Huang (2014) note how active the Chinese-Timorese are on Facebook, having established many Facebook groups and using this platform to organize get-togethers and cultural events.

4   Timor Hakka[16]

Berlie (2015: 38) describes the Chinese-Timorese community as being characterized by a "Hakka-dominant homogeneity". This description is apt synchronically, but, as we have seen above, historically, speakers of two distinct Hakka dialects as well as Hokkien and Cantonese were present in East Timor. Hakka speakers started to arrive in large numbers in the second half of the nineteenth century, and by the early twentieth century, they made up

---

16   Many thanks to my Timor Hakka consultants, and especially Mr. Lay Sum Li, for patiently answering my many questions.

the majority of the Chinese-Timorese community. In 1975, Kammen & Chen (2019: 79) estimate that between 85–90% of the community were Hakka speakers. Today, few Cantonese speakers, most of whom are elderly, remain in Timor-Leste. Their descendants have mostly switched to Hakka (Lay Sum Li p.c., 16 March 2020). My consultants are not aware of any Hokkien speakers among them (cf. Table 2.2).

The Hakka variety of East Timor is referred to here as Timor Hakka. Not many people in Timor-Leste speak Timor Hakka today: according to the 2015 population census (General Directorate of Statistics 2015) the number of mother tongue speakers of Chinese in the country is 827, of whom 599 live in the capital Dili. The next highest concentrations, with 51 and 42 speakers respectively, are reported for the Liquiça and Manatuto districts, which abut on the Dili district to the east and the west along the north coast of Timor. It is not clear, however, that the term "Chinese" used in the census publications refers to Timor Hakka. The number of Chinese nationals living in Timor-Leste is given in the census as 850. The near-coincidence of these figures suggests that the number of 827 Chinese mother tongue speakers refers to recent immigrants from China.[17] If this interpretation is correct, Timor Hakka is not included in the census at all. In that case, its speakers might be subsumed in the "Other" category, which numbers 617 speakers. It is thus hard to make an accurate statement on the number of Timor Hakka speakers in Timor-Leste today.

As noted in the previous section, Soares (2019: 317) estimates that some 4,000 Chinese-Timorese lived in Timor-Leste in 2014. Assuming the 827 mother tongue speakers of Chinese reported in the 2015 census are speakers of Timor Hakka, the discrepancy between these numbers seems to suggest that as few as one-fifth of people who identify as Hakka or Chinese-Timorese actually speak Hakka as a mother tongue. Indeed, it is not uncommon to meet people who have Hakka surnames and identify as Hakka, but do not speak the language. It is my impression that these people are predominantly in their thirties or younger. Likewise, in Australia, language competence in Timor Hakka

---

17   The same unclarity is found in volume 2 of the publications accompanying the 2010 census (General Directorate of Statistics 2010), which reports a number of 722 mother tongue speakers of Chinese living in Timor-Leste on p. 204. P. 196 of the same report gives the number of Chinese nationals living in the country as 1,139. On p. 11, the report states that "[f]or non-Timorese people, the person's main language was recorded", which would suggest that the 722 Chinese speakers include also post-independence newcomers. In this case, however, the discrepancy between the number of Chinese nationals, who surely are native speakers of a Sinitic language, and that of mother tongue speakers of Chinese would require explanation.

is decreasing in the second and following generations, following the typical pattern of three-generation shift to English (Chew & Huang 2014: 314–16). As discussed in section 2.2, Hakka does not have official recognition in Timor-Leste. It is primarily a home language and is limited to the oral domain. No Timor Hakka literature is known, and whether there are distinct stylistic repertoires is unclear. It is not generally learned by community outsiders (although indigenous East Timorese employees in Hakka-owned family businesses may pick it up and speak it well if they work with the family in close proximity; Lay Sum Li p.c., 21 August 2017).[18] All speakers are at least bilingual with Tetun (Timor-Leste's Austronesian lingua franca), and depending on their biography, may know a variety of additional languages such as Portuguese, Indonesian, Mandarin, English, or other Timorese languages. In sum, with a low number of speakers below the age of 30, no official recognition, and a limited domain of use, Timor Hakka is an endangered language variety.

To date, no information on the linguistic characteristics of Timor Hakka is available. Anecdotal evidence suggests that a distinct localized Timorese variety of Hakka has developed in East Timor: according to John Hajek (p.c., 20 February 2015), when Australian authorities sent Hakka speakers from Malaysia to act as interpreters for Chinese-Timorese immigrants, they had trouble communicating; and one of Wise's (2006: 149) interviewees says: "Our Hakka has evolved and is different to mainland Chinese Hakka. Sometimes it's a problem, because it means we tend to mix more with only Chinese-Timorese", likewise suggesting that Timor Hakka speakers may not be able to communicate easily with speakers of other Hakka varieties. Speakers of Timor Hakka in Dili and more recent Chinese-Indonesian immigrants from Kalimantan, on the other hand, acknowledge lexical differences between their respective Hakka varieties, but can communicate without major problems. In any case, Hakka is an important part of the Chinese-Timorese identity (Wise 2006: 149).

A number of factors may have contributed to the emergence of Timor Hakka as a distinct variety differing from those spoken in e.g. Malaysia or Indonesia: a) the particular Hakka dialect brought to Timor from China by a significant portion of the original immigrants; b) the development of a koine variety in Timor by immigrants speaking different dialects; and c) contact with local

---

18   Some degree of integration of local people into Hakka society appears to have been fairly common in Hakka settlements in rural Indonesia: according to Heidhues (1996: 177), local wives and their children commonly adopted the Hakka language and Chinese dress. Thanks to Tom Hoogervorst for making this source available to me.

languages as well as the colonial language.[19] Kammen & Chen (2019: 79) show that the Chinese-Timorese community is predominantly made up of Meixian Hakka speakers (cf. Chew & Huang 2014: 306). Table 2.3 provides a preliminary lexical comparison of Timor Hakka with the Meixian dialect (data from Hashimoto 2010) and the variety spoken around Hong Kong (Bao'an 寶安; data from Chappell & Lamarre 2005).[20] Direct comparison is complicated somewhat due to the use of different romanization systems in these sources: Chappell & Lamarre's (2005) description of Hong Kong Hakka, an annotated edition of late nineteenth- and early twentieth-century materials from the Basel Mission to China, follows the original materials in using the Lepsius script. For Meixian Hakka, Hashimoto (2010), originally published in 1973, uses an IPA transcription expanded by some non-standard symbols from the sinological tradition.[21] Timor Hakka forms come from my own data and are in a broad phonetic transcription; a full phonological analysis remains to be done. In Table 2.3, I have attempted to convert the data into a uniform, broad phonetic romanization, which is given in the first line of every cell. The data from Meixian Hakka and Hong Kong Hakka include a second line, which reproduces the romanization as found in the original source.[22] Tones are omitted, as the tonology of Timor Hakka has not yet been analysed.

Table 2.3 shows that overall, Timor Hakka seems to align more with Meixian Hakka than with Hong Kong Hakka. There are, however, some cases where Timor Hakka forms are more similar to those recorded for Hakka as spoken in the Hong Kong region; a case in point is Timor Hakka *swi* 'water'. There is also a non-negligible number of cases where Timor Hakka forms correspond fully to neither dialect, displaying either other dialectal forms or innovative forms.

---

19   Cf. Skinner (1967: 104–5), according to whom the Hakka variety spoken in north-eastern Bangka (Indonesia), is "almost creolized … with heavy borrowings from Bangkanese Malay". Thanks to Caroline Chia for bringing this source to my attention.

20   I focus in this section on the lexicon, as I have as yet done little analysis of Timor Hakka. Language change may of course affect all linguistic levels, from (segmental as well as suprasegmental) phonology to morphology and syntax.

21   An additional complexity of this source is that different conventions are used on different analytical levels: a narrow phonological transcription is used in the syllabary in the phonetics section, a somewhat broader transcription in the remainder of the phonetics section, and a romanization based on a phonological analysis in the remainder of the book. For instance, [ŋ] and [ɲ] are analysed as allophones. Thus, outside the phonology section, both are represented by ⟨ŋ⟩; and [ʃ], which is an allophone of both /s/ and /h/, is represented as either ⟨s⟩ or ⟨h⟩ (Hashimoto 2010: 101–2).

22   All errors in the conversion are my own.

TABLE 2.3   A lexical comparison of Meixian Hakka, Hong Kong Hakka, and Timor Hakka

English	Meixian Hakka	Hong Kong Hakka	Timor Hakka
(durative marker)	[ten] (den)	[kin] (kin)	[teŋ]
(genitive marker)	[e] (/e/)ᵃ	[kaj] (kai)	[ke] ~ [ge] ~ [e]
1PL(EXC)	_ᵇ	[ŋa-teu] (nga teu)	[ŋai-ten]
1PL(INC)	[ŋaj-ten-ɲin] (/ŋai-den-ŋin/)	[ŋaj-teu] (ngai teu)	[tɕi-ga]
banana	[kiuŋ-tsiau] ([kjiʊŋ-tsiɑʊ])	[ŋa-tsjaw] (nga tsyau)	[tjuŋ-tɕeu]
blood	[çjat] (/hiat/, [ʃiæt])	[çjet] (hyet)	[sjat]
dog	[kew-e] ([gɛu vɛ])ᶜ	[kew] (keu)	[kew-li] ~ [kjew-li]
five	[ŋ] (/ŋ/)	[ŋ] (ng)	[m]
sky	[tʰjen-ɲi] (/tien-ŋi/)	[tʰen] (then)	[tjen]
ten	[sɤp] ([sɿp])	[ʃip] (šip)	[sup]
this	[li] (/li/)	[ɲa] ~ [le] (nya ~ le)	[ti]
water	[se] (/se/)	[ʃwi] (šui)	[swi]

a   The marker appears to assimilate to the preceding noun or pronoun, e.g. *ŋai-je* 'my', *ŋ-ŋe* 'your' etc. (e.g. Hashimoto 2010: 468).
b   According to Hashimoto (2010: 451), there is no clusivity distinction in the Meixian dialect. However, Lau (2017) describes a distinction for the nearby Dabu 大埔 dialect, giving the first person plural inclusive form as *ɛn* (*tɛu*) and the first person plural exclusive as *ŋai tɛu*.
c   Lau (2017).

A conspicuous feature of Timor Hakka is the relatively widespread use of the suffix *-li*, especially with names of animals, but also in other semantic domains; *keu-li ~ kjew-li* 'dog', as in Table 2.3, is an example; (1) gives a few additional ones. This is reminiscent of the diminutive suffix *-tsai* as described by Lau (2017) for Hong Kong Hakka.

(1)   *ma-li*   'horse'   *ke-li*   'rabbit'
       ɲu-li   'cow'   *tɕu-li*   'pig'
       fa-li   'flower'   *tʃʰa-li*   'car'
       laj-li   'man'   *moj-li*   'woman'
       to-li   'knife'   *ko-li*   'song'

Further study is necessary to determine whether the lexicon of Timor Hakka matches a particular Hakka dialect other than those shown in Table 2.3, or whether it combines mixed dialectal forms. A mix of dialectal features in Timor Hakka would point to koineization, and given the diverse linguistic makeup of immigrants as discussed above, it is likely that this process made at least some contribution to the emergence of Timor Hakka. A fact worth mentioning here is that my pilot fieldwork in Timor-Leste in 2017 with two speakers revealed noticeable heterogeneity in those language samples. The examples in Table 2.4 show that variation is found both in loanwords as well as Sinitic vocabulary; it concerns both content words as well as discourse markers, and there are both what appear to be pronunciation variants as well as etymologically unrelated lexemes. Speaker A is a man aged around 40, whose ancestors had come to Timor three or four generations ago and were likely Meixian Hakka. Speaker B is a man in his fifties, whose family has lived in Timor for 7 generations. He reports that his ancestors were brought to Timor from Macao as convicts.

Research into the factors determining these variants is yet to be carried out. In the meantime, a variety of factors can be hypothesized to be relevant. An important factor relates to the linguistic background of a speaker's family. For instance, the descendants of Cantonese convicts may have retained

TABLE 2.4   Lexical variation in Timor Hakka

English gloss	Speaker A	Speaker B
bus	*bas* (< English bus)	*masibombu* (< Portuguese *machimbombo*[a])
cat	*mjan-kuŋ*	*meu-kuŋ*
child	*se-ɲin*	*se-kwei*
enter	*ɲip*	*ɲup*
then; afterwards	*jang-heu*	*kjak-mi*

a   According to Pons Online Portuguese-English dictionary (https://en.pons.com/translate/portuguese-german/machimbombo#dict), *maximbombo* 'bus' is characteristic of Angolan and Mozambique Portuguese. It is thought to be ultimately derived from the English phrase *machine pump*, a term used to refer to Lisbon's mechanic lifts (https://ciberduvidas.iscte-iul.pt/consultorio/perguntas/a-origem-da-palavra-machimbombo/24503).

different linguistic features from descendants of Hakka-speaking immigrants; and the speech of those whose families originated in the Meixian area may show different dialectal features than that of other Hakka immigrants. It has also been shown in section 2 that linguistic background often correlates with other historical and demographic factors which are potentially relevant, that is the time of immigration and the likelihood of intermarriage with indigenous Timorese. The more recent the immigration event, the more likely a family's lect is to resemble the dialect of their mainland Chinese place of origin. On the other hand, the longer a family has lived in East Timor, the more likely it may be to see language contact effects. The majority of Chinese immigrants in the twentieth century were Meixian Hakka, while the descendants of Hokkien-speaking families have lived in East Timor for a considerably longer time. Cantonese speakers tended to be shunned by Hokkien and Hakka speakers and be of a lower socio-economic status. They were more likely to marry East Timorese women, and thus their descendants would have grown up in mixed households. This in turn may have resulted in a stronger linguistic influence of East Timorese languages. The same is true for families residing in rural areas, where often there were few other ethnic Chinese, or none at all. In Dili, on the other hand, there was a proper Chinese community and there may have been a stronger normative pressure acting on linguistic forms.

A further possible factor might be a speaker's age. East Timor's turbulent history is very directly reflected, for instance, in speakers' education. Older speakers who received the bulk of their education before 1975 would have gone either to a Portuguese or a Chinese school. As a result, Portuguese may be an obvious donor language for these speakers when it comes to linguistic borrowing. For those who went to school between 1975 and 2002, it would be Indonesian, as this was the sole language of education at the time. Finally, younger speakers, who received their formal education in Timor after independence in 2002, were educated in a mix of languages: the official languages at school have since then been Tetun and Portuguese. However, many teachers kept using Indonesian for a considerable time after independence due to the lack of Tetun-language teaching materials and their own insufficient fluency in Portuguese. Another factor related to an individual speaker's biography is whether he or she spent a considerable period abroad, and if so where. As mentioned, many Chinese-Timorese left Timor-Leste during the Indonesian rule, emigrating either to Australia, Macao, or Indonesian West Timor. Again, this had obvious consequences for their linguistic inventory and may influence their Hakka. A case in point are the two terms used in Timor Hakka for 'bus' as seen in Table 2.4: speaker A used the English loanword *bas* 'bus'. As an adolescent, he left Timor as a refugee and lived in Australia for several years.

Speaker B, who did not share this background, used *masibombu* 'bus'. When asked about this term, speaker A acknowledged knowing it as "an old word for 'bus'". Finally, it is also possible that a speaker's gender influences his or her Hakka variety, at least in middle-aged and older speakers: women of those age groups did not leave the house very often in their youths, whereas younger women enjoy considerably more freedom. As a consequence, middle-aged and older women may be less fluent in other languages than men of the same age.

Language contact and lexical borrowing have already briefly been mentioned, and it is obvious that language contact contributed to the unique character of Timor Hakka. Numerous loanwords are found from Portuguese, as in (2), but also from local languages (3). Given that I have a limited amount of data and have worked with only two speakers, it is difficult at this stage to distinguish between established borrowings (used widely throughout the speech community) and nonce borrowings, where an item from another language is spontaneously used by a bilingual speaker but may not occur in another instance or be used by other speakers. As an approximation, I have classified the items in (4) as nonce borrowings because the speakers who used them were readily able to come up with 'real Hakka' translations when I inquired. I treat the items in (2) and (3) as established borrowings because they were either used by both speakers or recognized by both speakers, and no Hakka equivalent was readily produced.

(2) *makaku*   'car jack'   > Portuguese *macaco*, Tetun *makaku*
     kapoŋ   'bonnet, hood'   > Portuguese *capô*, Tetun *kapó*
     perekasa   'ferry, boat'   > Portuguese *barcaça* 'barge', Tetun *barkasa*

(3) *pasat*   'market'   > Indonesian *pasar*, Tetun *basar*[23]
     kopi tʃʰa   'coffee' (lit. 'coffee tea')   > Indonesian *kopi*
     kaŋko   'water spinach'   > Tetun *kanko ~ kanku*
                   (*Ipomea aquatica*)
     kulu   'jackfruit'   > Tetun *kulu*
                   (*Artocarpus heterophyllus*)

(4) *niki*   'bat'   > Tetun *niki*
     lutu   'fence'   > Tetun *lutu*
     maki   'angry'   > Indonesian *maki*

---

23 This loan is found, in the same form, also in Indonesian, Malaysian and Singaporean Hakka. Thanks to Tom Hoogervorst for commenting on this.

The linguistic character of Timor Hakka thus illustrates the point that the Sinophone "inevitably registers the multiple tongues spoken in constant inter-action and creolization with indigenous and other local languages in a given place" (Shih 2013: 8–9). The preliminary insights on Timor Hakka in the pre-ceding paragraphs have also shown how rudimentary our knowledge of this Hakka variety is. Much work thus remains to be done, and given the variety's endangered status, this is timely. The Hakka community of Timor-Leste as well as the expat community of Australia are well aware of the ongoing language loss and support efforts to document and describe their language.

5   Concluding Remarks

Ethnic Chinese people have been an established part of East Timor's scene for centuries. However, the Chinese-Timorese community has long been neglected in studies on Chinese-descended communities; only recently have a handful of publications on them appeared, most of which are historical in nature. It is one of the goals of this chapter to introduce the community to a wider readership within the context of Chinese-descended communities else-where in Southeast Asia.

The Chinese-Timorese were closely associated with the colonizers and have developed mixed cultural forms influenced by the long Portuguese rule. On the other hand, they have also lived side-by-side with the indigenous East Timorese largely peacefully, integrating into East Timorese society to varying degrees. Both in the colonial and post-independence periods, they have contributed significantly to building East Timor's economy and dominate the retail busi-ness until today. By the end of the colonial period in 1975, there was a small but thriving community of some 8,000 people, making up around 1.5% of the total population (although estimates vary; see section 3.2). This situation was brought to an abrupt end by East Timor's declaration of independence and the subsequent Indonesian invasion, a pivotal event in Timor-Leste's recent his-tory. It resulted in a 24-year occupation which brought much suffering not only to the Chinese-Timorese, but the whole population of East Timor.

At the very periphery of the Nanyang region, the Chinese-Timorese com-munity originally arose as a satellite of the Chinese community of Kupang, West Timor. By the early twentieth century at the latest, when the colonial spheres of influence on Timor had fully consolidated, the Chinese residents of East Timor can be considered a "Luso-Chinese" community separate from the "Sino-Dutch" one of West Timor (Pinto 2014: 163). The circumstances of the ethnic Chinese in East Timor can be assumed to have differed from

those in Netherlands Indies in a variety of respects. For instance, the long period spent under Portuguese colonial rule in East Timor left its mark in the form of both cultural and linguistic influences. It is known that East Timorese Chinese merchants had business relations with Chinese businessmen in ports in the Netherlands Indies such as Makassar and Surabaya as well as Singapore (Kammen & Chen 2019: 83, fn. 4, 93). However, it is likely that they had to overcome a variety of administrative and political hurdles in order to maintain them, whereas ethnic Chinese residents of the Netherlands Indies would have faced few obstacles in conducting business with each other. On the other hand, Portuguese Timor had closer links to Macao than the ethnic Chinese in the Netherlands Indies had. Berlie (2015: 38) furthermore notes that, in contrast to Chinese-Indonesians, who speak a range of Sinitic languages, the Chinese-Timorese are exclusively Hakka-speaking. Given the difference in size between Indonesia and Timor-Leste, it is of course hardly surprising that the ethnic Chinese minority of Indonesia is much more diverse. We have also seen that the Hakka-speaking homogeneity of the Chinese-Timorese community is relatively recent. Another contrast that has been noted in the literature is that attitudes toward the Chinese in Indonesia have been characterized as a "mixture of resentment over riches and outright racism" (England 1999). In East Timor, there are no reports of anti-Chinese excesses by indigenous Timorese. While there is some resentment, several sources stress that the Chinese-Timorese are an accepted part of Timorese society.

On the other hand, Kiernan (2003) points out that the recent history of the Chinese-Timorese, which was shaped by the Indonesian invasion, parallels in a variety of respects that of the Chinese community of Cambodia, which is discussed in McFarland (this volume). In 1990, Kiernan wrote that "[t]he Chinese under Pol Pot's regime suffered the worst disaster ever to befall any ethnic Chinese community in Southeast Asia" (Kiernan 1990). While the death toll among the Chinese-Timorese during the Indonesian invasion was lower, this event was probably equally disruptive. As a result of the events surrounding the Indonesian invasion, an estimated 4,000 Chinese-Timorese remain in Timor-Leste today, accounting for less than 0.4% of Timor-Leste's population. The majority of the community live in Australia. Wise (2006) as well as Chew & Huang (2014) show that the Chinese-Timorese have a strong sense of identity and undertake significant efforts to maintain community relations, preferring to socialize with other Chinese-Timorese rather than with the wider Timorese or Chinese expat communities.

Language is an important part of many communities' identity. The Chinese-Timorese speak a Hakka variety, which Wise (2006) confirms is a very significant identity marker for the expat community in Australia. Timor Hakka

is now acutely endangered, both in Timor-Leste as well as in the Australian diaspora. My fieldwork in Timor-Leste has shown that speakers are aware of and concerned about this situation. To date, little is known about the linguistic characteristics of Timor Hakka. A preliminary lexical comparison suggests it is similar to the Meixian dialect of Hakka as described in Hashimoto (2010), although there are also lexical items that differ. Whether Timor Hakka aligns with another Hakka dialect from the Guangdong province, or whether it is a koine which evolved in Timor requires further research. First field data have furthermore shown that there appears to be a significant degree of variation within Timor-Hakka. The factors determining the occurrence of different variants are still unknown; one possibility is that they reflect differences between the varieties or languages of the original immigrants. Finally, the effects of contact with the other languages present in Timor, including the colonial language Portuguese as well as local languages and lingua francas, is obvious. However, not only Timor Hakka, but Southeast Asian Chinese varieties in general are to this day a largely untapped resource in Western linguistics, despite their obvious potential to inform scholarship in language contact, language attrition and maintenance, and sociolinguistics.

References

Babo Soares, Dionisio, 'Political developments leading to the referendum', in James J. Fox and Dionisio Babo Soares (eds.), *Out of the Ashes: Destruction and Reconstruction of East Timor* (Canberra: ANU E Press, 2003): 53–73. https://press.anu.edu.au/publications/out-ashes.

Baxter, Alan N. and Hugo C. Cardoso, 'Early Notices Regarding Creole Portuguese in Former Portuguese Timor', *Journal of Language Contact*, 10/2 (2017): 264–317.

Berlie, Jean A., 'Chinese in East Timor: Identity, society and economy', *HumaNetten*, 35 (2015): 37–49.

Bonacich, Edna, 'A theory of middleman minorities', *American Sociological Review*, 38 (1973): 583–94.

CartoGIS Services, 'The position of Timor-Leste in relation to its closest neighbours', Canberra: College of Asia and the Pacific, The Australian National University (accessed 29 July 2019). http://asiapacific.anu.edu.au/mapsonline/.

Chappell, Hilary, 'Linguistic areas in China for differential object marking, passive, and comparative constructions', in Hilary Chappell (ed.), *Diversity in Sinitic languages* (Oxford: Oxford University Press, 2015): 13–52.

Chappell, Hilary and Christine Lamarre, *A Grammar and Lexicon of Hakka: Historical Materials from the Basel Mission Library* (Paris: École des Hautes Études en Sciences Sociales, 2005).

Chew, Daniel and J. Sonia Huang, 'The Timorese Hakka in Australia: Community and the Internet', *Global Hakka Studies,* 2 (2014): 301–36.

Coates, Timothy J., 'The Portuguese Empire, 1100–1932', in Clare Anderson (ed.), *A Global History of Convicts and Penal Colonies* (London: Bloomsbury Academic, 2018): 27–64.

Constable, Nicole, 'Introduction: What does it mean to be Hakka?', in Nicole Constable (ed.), *Guest People: Hakka Identity in China and Abroad* (Seattle: University of Washington Press, 1996): 3–35.

Department of Immigration and Citizenship, Australian Government, *Community Information Summary: Timor-Leste-born* (Canberra: Community Relation Section, Department of Immigration and Citizenship, 2016) https://www.homeaffairs.gov.au/about-us/our-portfolios/multicultural-affairs/community-information-summaries/country-list-of-summaries.

Disney, Anthony R., *A History of Portugal and the Portuguese Empire*, vol. 1. (Cambridge: Cambridge University Press, 2009).

Douglas, Karen Manges and Rogelio Sáenz, 'Middleman Minorities', in William Darity (ed.), *International Encyclopedia of the Social Sciences* (2nd edn. Detroit, Mich.: The Gale Group, 2007): 147–48.

Dunn, James, *Timor: A People Betrayed* (Milton, Queensland: The Jacaranda Press, 1983).

Durand, Frédéric, 'Three centuries of violence and struggle in East Timor (1726–2008)', *Online Encyclopedia of Mass Violence* (2011). https://www.sciencespo.fr/mass-violence-war-massacre-resistance/en/document/three-centuries-violence-and-struggle-east-timor-1726-2008 (8 August, 2019).

Eberhard, David M., Gary F. Simons and Charles D. Fennig (eds.), *Ethnologue: Languages of the World* (22nd edn. Dallas, Texas: SIL International, 2019). http://www.ethnologue.com.

Edwards, Owen and UBB, 'Languages of Timor', Kupang: Language & Culture Unit (UBB) and Leiden University (2018). https://vici.marianklamer.org/linguistic-maps.html (26 July, 2019).

England, Vaudine, 'Chinese legacy of fear in Dili', *South China Morning Post* (1999). Hong Kong. https://etan.org/et99b/september/1-4/1chinese.htm (6 August, 2019).

England, Vaudine, 'Catch a red taxi to downtown Dili', *South China Morning Post.* Hong Kong (2001). https://www.scmp.com/article/359255/catch-red-taxi-downtown-dili (6 August, 2019).

Fat, Paul B. Tjon Sie, 'They Might as Well Be Speaking Chinese: The Changing Chinese Linguistic Situation in Suriname under New Migration', in Eithne B. Carlin, Isabelle Léglise, Bettina Migge and Paul B. Tjon Sie Fat (eds.), *In and Out of Suriname: Language, Mobility and Identity* (Leiden: Brill, 2015): 196–228.

Figueiredo, Fernando Augusto de, *Timor: A presença portuguesa (1769–1945)* (PhD thesis. Porto: Universidade de Porto, 2004).

Flynn, Dennis and Arturo Giraldez, 'Silk for Silver: Manila-Macao Trade in the 17th Century', *Philippine Studies* 44/1 (1996): 52–68.

General Directorate of Statistics. '2010 Census publications', 2010. http://www.statistics.gov.tl/category/publications/census-publications/2010/ (12 September, 2019).

General Directorate of Statistics. '2015 Census publications', 2015. http://www.statistics.gov.tl/category/publications/census-publications/2015-census-publications/ (12 September, 2019).

General Directorate of Statistics. Statistics Timor-Leste. http://www.statistics.gov.tl/category/survey-indicators/timor-leste-in-figures/ (31 July, 2019).

Government of Timor-Leste. http://timor-leste.gov.tl/ (26 July, 2019).

Gunn, Geoffrey C., 'History of Timor', 1999. http://pascal.iseg.utl.pt/~cesa/History_of_Timor.pdf.

Gunn, Geoffrey C., 'The Timor-Macao Sandalwood Trade and the Asian Discovery of the Great South Land?', *Review of Culture* 53 (2016): 125–48.

Hägerdal, Hans, 'The Slaves of Timor: Life and Death on the Fringes of Early Colonial Society', *Itinerario* 34/2 (2010): 19–44. https://doi.org/10.1017/S0165115310000331.

Hägerdal, Hans, *Lords of the land, lords of the sea: Conflict and adaptation in early colonial Timor, 1600–1800* (Leiden: KITLV Press, 2012).

Hajek, John, 'Towards a language history of East Timor', *Quaderni del Dipartimento di Linguistica-Università di Firenze* 10 (2000): 213–27.

Hajek, John and Francesco Goglia, 'The East Timorese in Australia: Multilingual repertoires, language attitudes, practices and identity in the diaspora', *Journal of Multilingual and Multicultural Development* (2019): 1–18.

Hammarström, Harald, Robert Forkel and Martin Haspelmath, *Glottolog 4.0* (Jena: Max Planck Institute for the Science of Human History, 2019). http://glottolog.org (31 July, 2019).

Hashimoto, Mantaro J., *The Hakka dialect: A linguistic study of its phonology, syntax and lexicon* (2nd edn., Cambridge: Cambridge University Press, 2010).

Heidhues, Mary Somers, 'Chinese settlements in rural Southeast Asia: Unwritten histories', in Anthony Reid and Kristine Alilunas-Rodgers (eds.), *Sojourners and settlers: Histories of Southeast Asia and the Chinese* (1996): 164–82. Honolulu: University of Hawai'i Press.

Hirth, Friedrich and W.W. Rockhill, *Chau Ju-kua: his work on the Chinese and Arab trade in the twelfth and thirteenth centuries, entitled Chu-fan-chï* (St Petersburg: Printing

Office of the Imperial Academy of Sciences, 1911). https://archive.org/details/cu319 24023289345/page/n1/mode/2up (1 April, 2020).

Hull, Geoffrey, *The Languages of East Timor: Some Basic Facts* (Dili: Instituto Nacional de Linguística, Universidade Nacional Timor Lorosa'e, 2002).

Jannisa, Gudmund, *Timor-Leste in the World: BC to Independence* (Lund: A Malae Production, 2019).

Kammen, Douglas and Jonathan Chen, *Cina Timor: Baba, Hakka, and Cantonese in the Making of Timor-Leste* (New Haven: Yale University Southeast Asia Studies, 2019).

Kiernan, Ben, 'The Survival of Cambodia's Ethnic Minorities', *Cultural Survival Quarterly Magazine* (1990). https://www.culturalsurvival.org/publications/cultural-survival-quarterly/survival-cambodias-ethnic-minorities (31 October, 2019).

Kiernan, Ben, 'War, Genocide and Resistance in East Timor, 1975–99: Comparative Reflections on Cambodia', in Mark Seldon and Alvin Y. So (eds.), *War and State Terrorism: The United States, Japan, and the Asia-Pacific in the Long Twentieth Century* (Lanham, MD: Rowman & Littlefield Publishers, 2003): 200–33.

KPP HAM, Investigative Commission into Human Rights Violations in East Timor, 'Full Report of the Investigative Commission into Human Rights Violations in East Timor', in Richard Tanter, Gerry van Klinken and Desmond Ball (eds.), *Masters of Terror: Indonesia's Military and Violence in East Timor* (Lanham: Rowman & Littlefield Publishers, 2006): 21–66.

Kwartanada, Didi, 'Middlemen minority in an isolated outpost: A preliminary study of the Chinese in East Timor to 1945', 2001. Manuscript. Online at https://independent.academia.edu/DidiKwartanada (24.8.2018).

Lau, Chun Fat, 'Hakka Dialect', in Rint Sybesma (ed.), *Encyclopedia of Chinese Language and Linguistics* (Leiden: Brill, 2017): 469–83.

Lefort, Julie, 'Chinese languages spoken in Mauritius: An overview', 2018. https://ro.uow.edu.au/mauritius50/papers/1/7/.

Leong, Sow-Theng, *Migration and Ethnicity in Chinese History: Hakkas, Pengmin and their Neighbours* (Stanford: Stanford University Press, 1997).

Lombard-Jourdan, Anne and Claudine Salmon, 'Les Chinois de Kupang (Timor), aux alentours de 1800', *Archipel* 56/1 (1998): 393–428.

Luo, Xianglin, 客家研究導論 *Introduction to Hakka Studies* (Guangdong: Shishan Library, 1933).

Martin, Ian and Alexander Mayer-Rieckh, 'The United Nations and East Timor: From Self-Determination to State-Building', *International Peacekeeping* 12/1 (2005): 125–45.

Miao, Mu, Dandan Yin and Xinxin Zhang, 东帝汶: 帝力关帝庙 *East Timor: Dili Guandi Temple*, 2015. http://www.huaxia.com/zhsx/gggl/yw/2015/08/4521464.html (27 April, 2020).

Moutinho, Isabel, 'Exile at the Edges of Empire: Contemporary Writing in Portuguese', in Michael Hanne (ed.), *Creativity in Exile* (Amsterdam/New York: Rodopi, 2004): 71–86.

Mulyanto, Randy, 'Chinese in East Timor: former Portuguese colony a model of integration by immigrants from China', *South China Morning Post*, 2019. https://www.scmp.com/lifestyle/family-relationships/article/3030631/chinese-east-timor-former-portuguese-colony-model (9 October, 2019).

Nicol, Bill, *Timor: A Nation Reborn* (Jakarta: Equinox Pub, 2002).

Norman, Jerry, *Chinese* (Cambridge: Cambridge University Press, 1988).

Norman, Jerry, 'The Chinese dialects: phonology', in Graham Thurgood and Randy J. LaPolla (eds.), *The Sino-Tibetan languages* (London: Routledge, 2003): 72–83.

Ooi, Keat Gin, 'Kapitan China System', in Keat Gin Ooi (ed.), *Southeast Asia: A historical encyclopedia, from Angkor Wat to East Timor* (Santa Barbara: Abc-Clio, 2004): 711.

Ormeling, Ferdinand Jan, *The Timor Problem: A Geographical Interpretation of an Underdeveloped Island* (Groningen: J.B. Wolters, 1956).

Pan, Lynn (ed.), *The encyclopedia of the Chinese overseas* (Singapore: Archipelago Press, 1998).

Pinto, Paulo Jorge de Sousa, 'Visitors and Settlers: Notes on Timor and the Chinese as Cultural and Economic Brokers (Sixteenth to Nineteenth Centuries)', *Journal of Asian History* 48/2 (2014): 139–64.

Pinto, Paulo Jorge de Sousa, 'Traders, middlemen, smugglers: The Chinese and the formation of colonial Timor (18th–19th centuries)', in José Vicente Serrão, Bárbara Direito, Eugénia Rodrigues and Susana Münch Miranda (eds.), *Property Rights, Land and Territory in the European Overseas Empires* (Lisbon: CEHC, ISCTE-IUL, 2015): 267–77.

Ptak, Roderich, 'O transporte do sândalo para Macau e para a China durante a dinastia Ming', *Revista de Cultura* 1 (1987): 36–45.

Ptak, Roderich, 'Trade between Macau and Southeast Asia in Ming Times: A Survey', *Monumenta Serica* 54 (2006): 465–89.

Purcell, Victor, *The Chinese in Southeast Asia* (2nd edn., Oxford: Oxford University Press, 1966).

Rae, Ian and Morgen Witzel, *The overseas Chinese of South East Asia: History, Culture, Business* (Basingstoke: Palgrave Macmillan, 2008).

Raynor, Sophie, 'Timor-Leste's forgotten Chinese', *South China Morning Post*, 2019. http://lowyinstitute.com.au/the-interpreter/timor-leste-forgotten-chinese (6 August, 2019).

Sagart, Laurent, 'On distinguishing Hakka and non-Hakka dialects', *Journal of Chinese linguistics* (1998): 281–302.

Sagart, Laurent and Hilary Chappell, 'Le hakka', in Emilio Bonvini, Joëlle Busuttil and Alain Peyraube (eds.), *Encyclopédie des Sciences du Langage: Dictionnaire des langues* (Paris: Presses Universitaire de France, 2011): 1016–24.

Scambary, James, 'Anatomy of a Conflict: The 2006–2007 Communal Violence in East Timor', *Conflict, Security & Development* 9/2 (2009): 265–88.

Skinner, G. William, 'The Chinese Minority', in Ruth T. McVey (ed.), *Indonesia* (New Haven, Connecticut: Yale University, 1967): 97–117.

Soares, Laurentina "Mica" Barreto, 'Overseas Chinese, soft power and China's people-to-people diplomacy in Timor-Leste', in Andrew McWilliam and Michael Leach (eds.), *Routledge Handbook of Contemporary Timor-Leste* (New York / London: Routledge, 2019): 314–30.

Sousa, Ivo Carneiro de, 'Para a História das relações entre Macau e Timor (séculos XVI–XX)', *Revista de Cultura Edição Internacional, Macau* 18 (2006): 13–22.

Subrahmanyam, Sanjay, *Portuguese Empire in Asia, 1500–1700: A Political and Economic History* (2nd edn., Hoboken: Wiley-Blackwell, 2012).

Tan, Chee-Beng (ed.), *Routledge handbook of the Chinese diaspora* (Oxon & New York: Routledge, 2013).

Tan, Chee-Beng, *Chinese Religion in Malaysia: Temples and Communities* (Leiden: Brill, 2018).

Telkamp, Gerard J., 'De Ekonomische Struktuur van Portuguees-Timor in de Twintigste Eeuw: Een Voorlopige Schets', Amsterdam, unpublished manuscript, 1975.

Telkamp, Gerard J., 'The Economic Structure of an Outpost in the Outer Islands in the Indonesian Archipelago: Portuguese Timor 1850–1975', in Francien van Anrooij (ed.), *Between People and Statistics: Essays on Modern Indonesian History Presented to P. Creutzberg* (Dordrecht: Springer Netherlands, 1979): 71–89.

Thaler, Kai, 'Foreshadowing Future Slaughter: From the Indonesian Killings of 1965–1966 to the 1974–1999 Genocide in East Timor', *Genocide Studies and Prevention*, 7/2–3 (2012): 204–22.

The Heritage Foundation, '2019 Index of Economic Freedom', 2019. https://www.heritage.org/index/country/timorleste (9 August, 2019).

Timor Tourism. 'The Chinese temple', *Timor Tourism*, 2015. https://web.archive.org/web/20151125045647/http://timor-tourism.tl/poi/religious-building/dili/chinese-temple (8 October, 2019).

Treaty of Lisbon, 'Tratado de demarcação e troca de algumas possessões portuguesas e neerlandesas no arquipélago de Solor e Timor, entre sua majestade El-Rei de Portugal e sua majestade El-Rei dos Países Baixos, assinado em Lisboa pelos respectivos plenipotenciários aos 20 de Abril de 1859. Imprensa Nacional', 1861. https://books.google.ch/books?id=4gMMAAAAYAAJ&printsec=frontcover&dq=tr

atado+de+lisboa+de+1859&hl=pt-BR&ei=3afNS_EFxPfwBq7PyWg&sa=X&oi=book
_result&ct=result&redir_esc=y#v=onepage&q&f=false (27 September, 2019).

UNESCO Bangkok: Asia and Pacific Regional Bureau for Education, *Overview of internal migration in Timor Leste*, 2018. http://data.unescap.org/escap_stat/#home (30 July, 2019).

United Nations, 'UNMIT: United Nations Mission in Timor-Leste', 2019. https://unmit .unmissions.org/ (9 August, 2019).

Wang, Gungwu, *Don't leave home: Migration and the Chinese* (Singapore: Eastern Universities Press, 2003).

Wise, Amanda, *Exile and return among the East Timorese* (Philadelphia: University of Pennsylvania Press, 2006).

Wurm, Stephen A., Rong Li and Theo Baumann, *Language atlas of China* (Australian Academy of the Humanities; Longman Group (Far East), 1987).

Language Contact and Lexical Changes in Khmer and Teochew in Cambodia and Beyond

Joanna Rose McFarland

1 Introduction[1]

Chinese people have formed vibrant communities throughout Southeast Asia and around the world, in which their culture and language are still very much alive today. One example of this is the Teochew people, who originated from the Chaoshan (潮汕; Chaozhou-Shantou) region of eastern Guangdong. From the eighteenth to the twentieth centuries, many Teochew people emigrated from China around Southeast Asia, including to Cambodia, Indonesia, Malaysia, Singapore, Thailand, and Vietnam. Teochew (潮州; also called Chaozhou, Chiuchow, Swatow, or Teochiu) is a member of the Sinitic Southern Min cluster and varieties in and out of China have been featured in several studies over the years. This research is explored below. Yet the available literature is still lacking. This is especially true for the Teochew of Cambodia, where the group makes up an abundant majority of the ethnic Chinese there, having resided in the country for several generations. Despite governmental oppression of these peoples during the 1970s and 1980s, many Cambodian Chinese continued speaking their native Sinitic languages and passed them on to the next generations, albeit with some changes. Continued contact with Khmer, the official language of Cambodia, has created an environment conducive to language change. Thus, when comparing Teochew varieties, we should expect Cambodian Teochew to have some differences in its sound system (phonology), grammar (syntax), and vocabulary (lexicon). This chapter

---

1 This chapter represents a portion of the researcher's work on creating a comprehensive reference grammar of Cambodian Teochew. For additional comparisons between Cambodian Teochew and the varieties spoken in Chaoshan and Southeast Asia, including phonological and grammatical divergences, see McFarland (forthcoming). Throughout this chapter, Chinese characters have been converted to simplified Chinese in accordance with contemporary Cambodian preferences, including those of the consultants in this study.

focuses on the lexical differences by highlighting Khmer loanwords that are unique to the vocabulary of the Cambodian Teochew variety.[2]

The earliest works on Teochew come from foreign missionaries to China and include dictionaries (Duffus 1883; Fielde 1883; Goddard 1888) and phrasebooks (Ashmore 1884; Dean 1841; Fielde 1878; Giles 1877; Lim 1886). In recent times, scholars have looked at Chaoshan varieties of Teochew from both phonetic/phonological (Bao 1999; Hong 2013; Li 2014; Lin 1995; Lin & Chen 2011; Lin 2019; Xu 2016) and syntactic (Matthews, Xu, & Yip 2005; Matthews & Yip 2008; Xu & Matthews 2011, 2013) viewpoints. For more comprehensive sources, Li (1959) and Xu (2007) provided full and partial, respectively, grammars of the Jieyang 揭阳 variety and highlighted some of the ways in which Teochew differs from Mandarin and other Sinitic languages. Lin Lunlun (1996) looked at the Chenghai 澄海 variety, also making comparisons to Mandarin and Middle Chinese.

Aside from varieties spoken in China, Low (2014) and Yeo (2011) presented grammatical sketches of Singapore Teochew that covered detailed phonology, the structure of noun and verb phrases, interrogative sentences, particles, and more. Low additionally provided a short list of Malay and English loanwords found in Singapore Teochew. While neither source offered a comprehensive grammar, both projects provided good basic data on the language variety. Goh (2017) created a dictionary of words and phrases used in everyday Singapore Teochew. Also included was a separate list of Singapore Teochew terms that have been borrowed from languages such as English and Malay. Li (1991) outlined the sound system of Singapore Teochew and presented a list of 100+ loanwords from English and Malay. Additionally, Li provided 50+ "special words" that arose from the influence from Hokkien and other Sinitic languages spoken in Singapore. From a formal syntactic standpoint, Cole and Lee (1997) looked at yes/no question formation while Cole Hermon, and Lee (2001) explored long distance reflexives in Singapore Teochew.

Next door, in Malaysia, Khoo Kiak Uie studied tonal differences (2017b), phonetic variation such as vowel nasalization and word onset alternations (2017a), and dialect mixing and a unique tone phenomenon (2018) of Teochew fishing villages in North Perak, Selangor, and Parit Jawa respectively.

On varieties spoken in Indonesia, Peng's (2012) doctoral dissertation studied syntactic constructions in Jambi Teochew and Pontianak Teochew that

---

2 While I use the term "Cambodian Teochew" to refer to the variety of Teochew spoken by Cambodian Teochew people, it is likely that there is variation amongst Teochew speakers in different regions of Cambodia, as Peng (2012) found in Indonesia with the Jambi and Pontianak varieties (see below).

differed from those in Jieyang, citing contact with Malay as a reason for the divergences. Focusing on relative clauses, Peng (2011) found that Jambi Teochew displayed both head-final (characteristic of Chinese languages) and head-initial (characteristic of Malay languages) word orders. Relative clauses also optionally borrow the Malay relativizer *yang*. Peng's (2012) main consultant for her Pontianak Teochew data was Yohana Veniranda, who herself studied Pontianak Teochew, specifically focusing on perfective aspect and negative markers and their interaction (Veniranda 2015), as well as acoustic differences between oral and nasal vowels and diphthongs (Veniranda 2016).

In Thailand, Atchariyasucha (1982) explored the phonology of Teochew in detail based on 1,280 words collected from one consultant, and made comparisons to the Thai sound system. Eiampailin (2004) worked with twelve native speakers of Teochew living in Bangkok who spoke Thai as a second language to study the phonological interference of Teochew on the speakers' Thai. Eiampailin's research included basic phonology of the target language based on the speech of the 12 consultants, which was largely identical to Atchariyasucha (1982) with some speaker variation in vowel nasalization. Other work includes that of Phadungsrisavas (2008), who compared the tone and tone sandhi systems of 20 speakers from five regions of Thailand, and Lin (2006), who examined Teochew loanwords in Thai and Thai loanwords in Teochew.

Chen (2009) explored the vernacular reading of Chinese texts by the Teochew in Cambodia, Laos, and Vietnam. The work included a brief outline of the phonology of each variety. Interestingly, the author noted the existence of initial /f/ in the Teochew of the Laotians and Vietnamese which I have not found written about in any other Teochew varieties.

As for Cambodia, in addition to Chen's (2009) introduction to the sounds of Cambodian Teochew, Pan (2000) showcased the pronunciation, vocabulary, and grammar of a third-generation Cambodian Teochew whose ancestors were from Jieyang. While the paper amounts to just eight pages and is only based on one speaker (with the pronunciation compared to one other), the research marks a good starting point for the documentation of Cambodian Teochew, even highlighting some divergences from other Teochew varieties such as vowel nasalization and the use of negation.

Overall, the body of research on Teochew is lacking, particularly for varieties outside of the Chaoshan region. This is unfortunate given the unique phenomena that were highlighted in the previous sources such as localized loanwords (Goh 2017; Li 1991; Lin 2006; Low 2014), and varietal phonological (Chen 2009; Khoo 2017a, 2017b, 2018; Pan 2000) and grammatical (Peng 2011, 2012; Pan 2000) divergences. This chapter, as well as McFarland (forthcoming), aims to bridge some of this gap by contributing novel work to the fields of Sinitic and

Cambodian linguistics, as well as to the field of contact linguistics, hopefully shedding light on the behaviour of languages in contact environments.

A few different romanization systems have historically been used for Teochew. The Pėh-oē-jī system, or "Church Romanization", was created in the nineteenth century by Western missionaries for writing all Sinitic Southern Min varieties. In 1960, the provincial government of Guangdong, China created a system known as the "Teochew Transliteration Scheme" or *Peng'im*, which is a Teochew transliteration of Mandarin *pīnyīn* 拼音. In 2002, language enthusiasts from the organization *Gaginang* 家己人 'our own people' began developing the Gaginang Peng-im System (GPIRS), which is an adaptation of Guangdong's Peng'im (TCKnow LLC. 2015). No speaker interviewed for this project wrote using one of these romanization schemes. Instead, when messaging friends or family, they would write in Khmer, Mandarin, or English, or send a voice recording in Teochew.[3]

Given my speakers' non-usage of romanization schemes, throughout this chapter, the collected data on Cambodian Teochew is presented in the International Phonetic Alphabet (IPA). Tone is indicated by numerical superscripts which represent the approximate pitch level of the voice on a five-point scale with [1] representing the low point and [5] the high point. When comparing data from the literature on Khmer and other Teochew varieties, words have been converted to the IPA for consistency. When not explicitly provided, tone pitch values have been approximated from tone markings in the original source.

The rest of the chapter is organized as follows: section 2 provides historical context for the Chinese in Cambodia before detailing their sociolinguistic situation as compared to Teochews in other parts of Southeast Asia, while also covering the Sinitic influence on Khmer, Thai, and Lao vocabulary; section 3 details the methodology for the present study and the findings and makes comparisons to other findings on loanwords in the world's languages; section 4 offers limitations to the current research, areas for future study, as well as some concluding remarks.

---

3   Low (2014) noted that the romanization schemes are also not used in Singapore. It has been shown that Indonesian Teochew people do transliterate Teochew in online spaces (see Birnie-Smith 2016).

2   Ethnic Chinese in Cambodia and Beyond

This section provides background information on the ethnic Chinese in Cambodia and their sociolinguistic situation and makes comparisons to Teochews in other parts of Southeast Asia. It also showcases some of the Chinese loanwords found in the languages of the region including Khmer and Thai.

2.1   *Note on Terminology*

There are several terms in the Khmer language to describe ethnic Chinese people.[4] Those who were born in China and emigrated to Cambodia are referred to as *cənchav* ចិនឆៅ, directly translated as 'raw Chinese'. Their descendants are known as *cən* ចិន 'Chinese', *kooncən* កូនចិន, literally 'children of Chinese' but also used to mean 'Cambodian Chinese', or *kooncavcən* កូនចៅចិន 'grandchildren of Chinese'. Those from mixed families can be described as *koonkatcən* កូនកាត់ចិន 'mixed Chinese', *koonkatkmae* កូនកាត់ខ្មែរ 'mixed Khmer', or *kmaekatcən* ខ្មែរកាត់ចិន 'mixed Khmer-Chinese', where *kat* translates to 'cut', suggesting a cut or split heritage. Recent Chinese immigrants from mainland China are known as *cəndaekook* ចិនដីគោក 'mainland Chinese'. Finally, *cəntiiəciiw* ចិនទាជីវ 'Teochew' refers specifically to those with Teochew heritage.

As for Teochew terms, *hua¹¹dziŋ²⁴* 华人 'Chinese' can be used for all ethnic Chinese and *hua¹¹khiau⁵⁵* 华侨 'Overseas Chinese' historically referred to Chinese citizens abroad. Nowadays it may be used for some speakers who do not have Chinese nationality. *Hua¹¹i²⁴* 华裔 'Overseas Chinese' is a term used for people of Chinese heritage with citizenship from another country. Most Cambodian Chinese are technically *hua¹¹i²⁴*, though my consultants reported unfamiliarity with the term, or rarely using it. The word *təŋ¹¹naŋ⁵⁵* 唐人 'Chinese', or people from the Tang 唐 Dynasty, is also used to mean 'Overseas Chinese'. Some Teochews in Cambodia use *təŋ¹¹naŋ⁵⁵* to refer specifically to Teochew people. This may be due to the group's current and historical majority among the country's ethnic Chinese population. Some of the nuanced differences between these terms appear to have been lost over the generations, which is why there are reports of the terms being used interchangeably. Finally, Teochew people call themselves *tiɔ¹¹tɕiu³³naŋ⁵⁵* 潮州人 'Teochew people' or *ka³³ki¹¹naŋ⁵⁵* 家己人 'our own people'. In Cambodia, they may also say they are *kaŋ²⁴pɔu³³tɕe¹¹təŋ¹¹naŋ⁵⁵* 柬埔寨唐人 'Cambodian Chinese' or *kaŋ²⁴pɔu³³tɕe¹¹hua¹¹dziŋ⁵⁵* 柬埔寨华人 'Cambodian Chinese'.

---

4   The terms that follow have been adapted from Edwards (2009) according to their usages by consultants interviewed for this project.

In the English literature on the Chinese in Cambodia, the terms "Sino-Khmer" and "Sino-Cambodian" are often used. Sino-Khmer traditionally refers to people of mixed Chinese and Khmer heritage (Dorais 1991: 553; Edwards 2009: 176; Tan 2006: 155–56), while Sino-Cambodian refers to a broader group of ethnic Chinese people with Cambodian citizenship (Verver 2012: 49). Speakers consulted for this project were generally unfamiliar with the Sino-terms. One speaker preferred to identify as purely Cambodian and would only say Teochew when specifically asked if they were ethnic Chinese, while other speakers and their families identify as "Cambodian Teochew". Therefore, in this chapter, I follow Verver (2012) in adopting the labels "Cambodian Chinese", "ethnic Chinese", or more specifically "Cambodian Teochew".

2.2   *History of the Chinese in Cambodia*

The Chinese have had a long history in Cambodia, with the earliest records coming from Chinese diplomat Zhou Daguan's 周达观 detailed account in 1296 of the Chinese who had settled in an area that is today's Siem Reap (Zhou 2007). Since then, a variety of Chinese groups have made Cambodia their home, including the Hokkien, whose communities date back to the 1400s (Chan 2005), the Hainanese, who began migrating in 1675 (ibid.), the Cantonese, for whom mass migration started in 1679 (Willmott 1967), and the Teochew, who had an established community in Kampot in the 1800s (Chan 2005) and opened a dialect association (会馆 *huiguan*) in Phnom Penh in 1814 (Chen 2015). Huge waves of Teochew immigration came in the 1930s to the 1960s, mostly from Jieyang 揭阳, Chaoyang 潮阳, and Puning 普宁, and by the 1960s, the Teochew people overtook the Cantonese to become the majority Chinese group in Cambodia (Willmott 1967). Some sources provided insight into the group's population numbers over time. According to a Portuguese visitor to Cambodia in 1609, it was reported that 3,000 of the 20,000 inhabitants (15%) of Phnom Penh were Chinese (Schliesinger 2011: 199). By 1897, 22,000 of the 50,000 inhabitants (44%) of Phnom Penh were Chinese (Siphat 2017: 185). In Cambodia as a whole, there were reportedly 106,764 Chinese in 1874, representing 11% of the country's population (Schliesinger 2011: 199). By 1967, the Chinese had grown to 425,000, making up 7.4% of the Cambodian population, with 75% reportedly of Teochew origin (Willmott 1967: 17).

Willmott's (1967) detailed ethnography reported that the Chinese were thriving in Cambodia in the 1960s. There were 200 Chinese schools across the country, various associations by regional group (Cantonese, Hakka, Hainanese, Hokkien, and Teochew), 31 sports clubs in Phnom Penh alone, five Chinese newspapers in the capital city, dedicated cemeteries, and a Chinese hospital (ibid. 87–89). Unfortunately, shortly after he completed his study, the Chinese

faced some troubling times in Cambodia. Oppression of the group began with the rule of Lon Nol from 1970–75 and his forced closings of Chinese language schools and newspapers, while the Khmer Rouge regime that lasted from 1975–79 led to the deaths of nearly one-half of the ethnic Chinese in Cambodia (Clayton 2006). Finally, when the Vietnamese took over in the 1980s, the Chinese faced continued suppression in retaliation for China's support of the Khmer Rouge (ibid.). As a result of this decades-long oppression, Cambodia saw a mass emigration of the Chinese population, most of whom would never return. Instead, many families settled in places such as Australia (Stevens 1990), Canada (Dorais 1991), France (Aw 2019; Panh & Bataille 2012; Tan 2006), and the United States (Ly 2000; Tan 2006; Ung 2000). By 1984, there were reportedly only 61,400 Cambodian Chinese in the country (Siphat 2017: 185). Of those who did stay, many experienced a loss of their Chinese language and culture such that "a generation whose parents were principally Teochew speakers ... grew up speaking Khmer" (Nyíri 2015: 15). Fortunately, some did practice their language in secret and were able to pass their Chinese varieties along to the next generations (Clayton 2006; Edwards 2009).

Recent times have seen a resurgence of the Chinese in Cambodia. They have re-established their strength in business (Siphat 2017; Verver 2012), with the help of new migrants from mainland China, Hong Kong, and Taiwan (Chin 2017). There has been a transformation of Chinese media with the establishment of five newspapers, two magazines, and one radio station (Nyíri 2015). Many of the old *huiguan*s have been resurrected (Chen 2015), and Chinese schools have reopened (Bourgerie 2017). Even still, it is challenging to calculate the number of ethnic Chinese in Cambodia because many remain fearful of admitting their background, though there are recent estimates. The Association of Chinese Nationals in Cambodia estimated the population of "pure Chinese" (i.e. not Sino-Khmer) to be between 300,000 and 340,000 in 1995 (Edwards 2009). The 2008 census put the number of Chinese at 1% of Cambodia's 15.5 million people (155,000) (Siphat 2017: 185), though that was likely an inaccurately low figure. In 2014, The Foundation of Associations of Chinese estimated the population to number around 1 million (Bourgerie 2017: 166). Chen (2009) and Edwards (2009) reported that 80% of the Chinese in Cambodia are of Teochew origin.

2.3 *Sociolinguistic Situation of the Teochew in Cambodia and Southeast Asia*

While many of the aforementioned sources provided good historical background on the different Chinese communities in Cambodia and information on the governmental persecution they have faced over the years, limited work

has been done on the language of the Cambodian Chinese, though some studies have provided insights into their sociolinguistic situation. Verver (2012) and Verver and Koning (2018) studied Chinese business practices in Phnom Penh and discovered that the majority of the entrepreneurs were Teochew families who had been in Cambodia for two to four generations. They and other scholars have found most of these families to be multilingual, maintaining their Sinitic variety like Teochew in addition to learning the national language of Khmer, with many, but not all, also studying Mandarin in schools (Chan 2005; Tea & Nov 2009; Verver 2012; Verver & Koning 2018; Willmott 1967). Bourgerie (2017) investigated this multilingualism, focusing on the educational system of the ethnic Chinese in Cambodia by collecting data from surveys on language use and language background in four major Chinese schools in Phnom Penh from 146 respondents in grades 11 and 12. The author found that the younger generation of Chinese are "proud of their heritage and see their identity as a status marker" (Bourgerie 2017: 177). Pride in a Teochew/Chinese heritage and the significance of a Chinese identity, especially for entrepreneurship, was also noted by Verver (2012) and Verver & Koning (2018). Meanwhile, it was found that language use and choice is situational. While Khmer, Mandarin, and other Sinitic varieties are all used to some extent at home (Bourgerie 2017; Nyíri 2012), Khmer is used with non-Chinese peers (Bourgerie 2017), Teochew (and Cantonese) is spoken with ethnic peers (Bourgerie 2017; Chen 2015) and often in business (Tea & Nov 2009; Verver & Koning 2018), and Mandarin is spoken in school and for association matters (Chen 2015: 121), and in business relations with mainland Chinese and in Southeast Asia (Verver & Koning 2018). Data from the current researcher's observations and interviews with speakers consulted for this project and their family members confirm these findings.

While Teochew is still being spoken to some extent by the younger generation in Cambodia, Khmer, Mandarin, and English are all playing a role in its decline. Similar language shifts are being seen in Indonesia, Malaysia, Singapore, Thailand, and more, for many of the same reasons. First, not speaking a minority language, such as Teochew, at home is detrimental to its transmission (Fishman 1991). It has been reported that many families have moved away from Teochew in the home domain in Cambodia (Nyíri 2012), Indonesia (Veniranda 2015), Malaysia (Ong 2018; Sim 2012; Wang 2016), Singapore (Lee 2015; Li et al. 1997; Ng 1996), and Thailand (Lee 2014). Of the families in Cambodia consulted for this study, it was similarly noted that children raised in households that primarily spoke Khmer were not likely to speak Teochew proficiently. Interestingly, in Thailand, Lee (2014: 191) found that some study participants were able to learn Teochew from business settings, despite speaking Thai at home, though this is very rare.

Institutional factors such as government policy and the language of education have also played a big role in each of these countries. As mentioned above, political policies during the 1970s and 1980s greatly crippled the Cambodian Teochew community's language use. Furthermore, Chinese schools in Cambodia began switching the language of instruction from regional varieties to Mandarin when the country established diplomatic relations with China in 1958 (Edwards 2009: 185; Goh 2011: 11; Willmott 1967: 88). When the schools reopened in the 1990s, instruction continued in Mandarin, resulting in the "spread of Mandarin among young people at the expense of Teochiu" (Nyíri 2012: 106). In Singapore, the launch of the governmental Speak Mandarin Campaign in 1979 was instrumental in a similar move away from regional varieties (Lee 2015; Li et al. 1997; Ng 1996). Teochew schools persisted with the support of clan associations until that time, at which point "cultural classes" were taught instead, typically in Mandarin (Li et al. 1997). Today, English is the language of instruction for all schools in Singapore, while Mandarin is offered as a subject. Neighbouring Malaysia had their own Speak Mandarin Campaign launched by the Selangor Chinese Chamber of Commerce in 1980, to great effect (Ong 2018; Sim 2012). Mandarin had already replaced local varieties in Chinese schools in the region, starting in the 1920s (Wang 2016).[5] Despite setbacks over the years – such as restrictions on government funding – at present there are over 1,000 Mandarin-medium schools across Malaysia (Ong 2018). National policy has also played a role in Indonesia where the use of Teochew took a big hit in the 1960s with the country's ban on Chinese culture and language (Stenberg 2015; Veniranda 2015), though the extent of the ban and its effects vary. While Veniranda wrote of personally experiencing discouragement from teachers of the use of Teochew and other mother tongues during her schooling in Pontianak in the 1970s and 1980s, two of her speaker consultants did not experience such a prohibition (Veniranda 2015: 20). Stenberg (2015) found that the relative remoteness and the density of Teochew communities in Pontianak helped stave off some of the effects of the ban, as compared to the Chinese in Java. Abolishment of the prohibition in 1999 led to a revitalization of Chinese culture and the use of Teochew, as well as the addition of Mandarin as an extracurricular activity or obligatory subject in schools (Veniranda 2015). In Thailand, schools began shifting to Mandarin instruction around 1930 (Chokkajitsumpun 1998). However, from 1939 to 1989, various governmental policies, nationalistic ideology, and pressures for the Chinese to excel in the Thai language and assimilate have restricted Mandarin education in different ways and reduced Teochew intergenerational transmission

---

5   Also see Hoogervorst (this volume) on the early history of Mandarin in Indonesia.

(Chokkajitsumpun 1998; Morita 2007). Since 1989, Mandarin has seen a revival due to changes in Thai national policy in support of the study of the language (Chokkajitsumpun 1998) and recent studies point to Mandarin being the most widely spoken Chinese language in Thailand today (Lee 2014; Rappa 2014).

Language practices are also shifting in business. Though Teochew is still dominant in the business community in Thailand (Lee 2014; Rappa 2014), and speaking the language is valuable in niche markets such as the rice trade across Thailand, Laos, Vietnam, and Cambodia, Mandarin has become the de facto lingua franca in entrepreneurship throughout Southeast Asia, along with English (Verver & Koning 2018). As such, Mandarin and English are seen as valuable languages to learn for business opportunities, resulting in an increase of speakers, often at the expense of Teochew as it is seen as a language of low prestige with little economic or instrumental value (Chokkajitsumpun 1998; Lee 2015; Li et al. 1997; Ng 1996; Quek 2013). For example, in the markets of Johor Baru, Malaysia, that used to be dominated by the Teochew language, Wang (2012) instead found Mandarin used in 90.9% of the business interactions.

In addition to economic reasons, a common trend found in the increasing use of Mandarin was the belief that speaking it would unite the various Chinese groups in the respective countries (Bourgerie 2017; Chokkajitsumpun 1998; Kyne 1999; Lee 2015; Li et al. 1997; Ong 2018; Sim 2012; Stenberg 2015). Finally, exogamy also lead to language shift for Teochews in Cambodia (Filippi 2010; Willmott 2006), Indonesia (Stenberg 2015; Veniranda 2015), Malaysia (Ong 2018; Sim 2012), Singapore (Ng 1996; Quek 2013), and Thailand (Lee 2014).

In summation, several factors are working against the continued transmission of Teochew throughout Southeast Asia. And it can be hard to reverse these trends. As Quek (2013) found, language shift in Singapore has led to a whole generation lacking the Teochew proficiency to pass the knowledge on to their children, even though they reported wanting to. Two of the families I worked with in Cambodia described similar situations and it is likely that this is also the case for many Teochews in Indonesia, Malaysia, and Thailand. Thus it is especially important to carry out language documentation work in these communities in order to create a lasting record of the languages. Additionally, working with communities and recording their speech is a way to further emphasize to them and others that their language is important and worthy of study, and to encourage intergenerational transmission.

2.4   *Sinitic Influence on Khmer and Thai*

Despite the more recent language shift, it has been shown that Cambodian Chinese have been (at least) bilingual for multiple generations. Thus, hundreds of years of contact between Khmer and Chinese has created an

environment conducive to language change. Though the focus of this chapter is on Khmer's influence on Cambodian Teochew, Chinese varieties have also been found to have affected Khmer vocabulary. This is a significant feature of the Sinophone, in which the local culture is in dialogue with other languages of that location (Shih 2013: 8). The impact of Teochew and other varieties is unsurprising given Cambodia's high percentage of Chinese historically, especially in regions such as the capital city of Phnom Penh. The group was and is "an aggressive social and economic minority" (Pou & Jenner 1973: 1), with significant importance and influence, especially in the business sector.

Pou and Jenner (1973) investigated Chinese lexical influence and found 300 words in Khmer that were borrowed from various Sinitic varieties including Cantonese, Hokkien, and Teochew. The authors attempted to attribute the source of each word to one of the Sinitic varieties. For example *cʰá* 炒 'fry'[6] that became Khmer *chaa* ឆា was attributed to Amoy Hokkien (ibid. 47), while *pʰ̆ue* 被 'blanket' that became Khmer *phuuaj* ភួយ was said to come from Teochew (ibid. 16). In several cases, the Chinese source was the same in multiple varieties, so the exact donor language could not be determined, such as with *só* 锁 'lock'. This became Khmer *saao* សោ, and was attributed to Amoy/Hokkien/Cantonese (ibid. 83). Pou and Jenner noted that the bulk of items appeared closest to Hokkien (1973: 4). This could be because the Hokkiens were reportedly the first Chinese group to settle in Cambodia (Chan 2005; Siphat 2017). The vocabulary items were typically new things introduced by the Chinese in Cambodia including terms from commerce and navigation (22%), food and articles of use (21% each), diversions such as gambling and theatre (6%), kinship terms (5%), technological terms (arts and crafts) (4%), administrative and legal (2%), religious terms (2%), miscellaneous verbs (8%), and miscellaneous nouns (7.5%). Huang and Mo (2017) also looked at Chinese loanwords in Khmer, claiming to have identified more than 250 borrowed Chinese words in Khmer, most which reportedly came from Teochew. As such, the authors compared the Teochew pronunciation to the Khmer loanword pronunciation and reported on any sound changes. Only 50 words were provided, many overlapping with Pou and Jenner (1973). It is unclear how many of Pou and Jenner's loanwords are still used in Khmer today and to what extent as several of the words they listed were not recognized by the *Vacanānukram Khmēr*

6   Words from Pou and Jenner (1973) are unchanged in form. The authors adopted a uniform system of transliteration for their Chinese data from various sources. Since they were unconcerned with tone, Chinese words appeared with the same diacritics as in their original source transcriptions.

dictionary and some were noted to be used only amongst Chinese speakers. According to Edwards (2012: 9), Chinese familial terms such as *jie* 姐 'big sister' and *yi* 姨 'auntie' are being used nowadays, even between Khmers.

Chinese has similarly had a great influence on other languages in Southeast Asia. Several works have looked at the language family's effect on the Thai language. Egerod (1958) presented 181 words from the Swatow 汕头 variety of Teochew that have been adopted into Thai, explaining how the words' pronunciations changed in the borrowed Thai forms. Similar to Pou and Jenner (1973), Egerod found that the Chinese loanwords were often related to commerce, kinship, food and beverage, and leisure events. Gong (2000) likewise looked at 315 Swatow loanwords in Thai, while Gyarunsut (1983) compiled a list of over 460 Chinese terms used in modern Thai. Cooper (2020) created an online dictionary of more than 750 Chinese roots, loans, and cognates in Thai using data compiled from Gyarunsut (1983), Gong (2000), and Manomaivibool (1975). Of these, nearly 350 are attributed to Teochew. Finally, Lin (2006), found 100+ Teochew loanwords in Thai in areas such as people designations, business activities, foodstuffs, and daily necessities.

Table 3.1 below compares person terminology borrowed from Teochew/ Chinese found in Thai, Khmer, and Lao. As mentioned above, exact language origins are difficult to pinpoint. The majority of the words in Table 3.1 were attributed to Teochew but a few were said to come from another Chinese variety.

TABLE 3.1   Teochew loanwords in Khmer, Thai, and Lao

Teochew	Khmer	Thai	Lao	English
(Ɂa^{33})hĭa^{33} (阿) 兄	(Ɂaa)hìiə ហ៊ី	(Ɂa:^{33})hia^{33} (อา)เฮีย		elder brother; Chinese male
(Ɂa^{33})sɔ^{55} (阿) 嫂	saao សោ	(Ɂa:^{33})sɔ:^{453} (อา)ซ้อ	(Ɂaa)sɔ̀ɔ	elder sister in law
(Ɂa^{33})sim^{51} (阿) 婶	(Ɂaa)sam សាំ	(Ɂa:^{33})sim^{453} (อา)ซิ้ม		aunt, wife of father's brother; Chinese woman
(Ɂa^{33})koŋ^{33} (阿) 公	(Ɂaa)koŋ កុង	(Ɂa:^{33})koŋ^{24} (อา)กง		grandfather[a]

a   Also Malay *əŋkoŋ* and Tagalog *iŋkoŋ* (Pou & Jenner 1973: 58).

TABLE 3.1   Teochew loanwords in Khmer, Thai, and Lao (*cont.*)

Teochew	Khmer	Thai	Lao	English
(ʔa³³)kow³³ (阿) 姑	kòow កូរ៉	(ʔa:³³)ko:³³ (อา)โก	(ʔaa)kōo	aunt, father's sister[b]
(ʔa³³)ku²⁴ (阿) 舅	(ʔaa)kuu[c]	(ʔa:³³)ku:²⁴ (อา)กู๋		uncle, mother's brother
(ʔa³³)mue²¹ (阿) 妹	(ʔaa)múuej មួយ	(ʔa:³³)muaj²⁴ (อา)หมวย	(ʔaa)mûej	younger sister; Chinese girl
(ʔa³³)pɛʔ²² (阿) 伯	(ʔaa)pèʔ ប៉ិ	(ʔa:³³)pɛʔ⁵⁵ (อา)แป๊ะ	ʔaapɛ̄ʔ	uncle, father's brother; old man
(ʔa³³)ti²⁴ (阿) 弟	(ʔaa)tii[d]	(ʔa:³³)ti:²⁴ (อา)ตี๋		younger brother
(ʔa³³)tsɛ⁵¹ (阿) 姐	caaɛ ចែ	(ʔa:³³)ʨe:⁴⁵³ (อา)เจ๊	cee	older sister
(ʔa³³)ʔi⁵⁵ (阿) 姨		(ʔa:³³)ʔi:⁴⁵³ (อา)อี๊		aunt, mother's sister
haj²⁴nam⁵⁵ 海南	hajnam ហៃណាម	haj²⁴lam²⁴ ไหหลำ		Hainan, Hainanese
hĩɔ³³koŋ³³ 香公		hia³³koŋ³³ เฮียกง		the keeper of a temple
huaŋ³³ 番		huan³³ ฮวน		foreigner; non-Chinese
kuŋ³³si³³ 军师		kun³³si:³³ กุนซือ		an advisor
now³³kĩa⁵¹ 孥仔		noŋ³³kiaʔ⁵⁵ โนงเกี๊ยะ		child
sã³³pũa⁵⁵ 三盘		sa:³³pua⁴⁵³ ซาปั๊ว		a third-level seller
sam³³pɔ²⁴koŋ³³ 三保公		sam³³pɔ:³³koŋ³³ ซำปอกง		a Chinese Buddha image

---

b   See also Malay *əŋkoh* (Hoogervorst, this volume Table 7.7).

c   This word is primarily used among Cambodian Chinese families. It is given in Huang & Mo (2017: 96) but is absent in Pou & Jenner (1973).

d   Primarily used among Cambodian Chinese families.

TABLE 3.1 Teochew loanwords in Khmer, Thai, and Lao (*cont.*)

Teochew	Khmer	Thai	Lao	English
si³³hĩa³³ 师兄		si:³³hia³³ ซือเฮีย		form of address for older male classmate
si³³tsɛ⁵¹ 师父		si:³³tɕe:⁴⁵³ ซือแป๋		form of address for teacher
sia¹¹ 舍		sia²² เสี่ย		a wealthy Chinese
siŋ³³sɛ̃³³ 先生	sinsaaɛ ស៊ិនសៃ	sin³³sɛ:²⁴ นแส		doctor; teacher
taj²¹koŋ³³ 舵公	tajkoŋ តៃកុង	taj⁴¹koŋ²⁴ ได้กั๋ง		helmsman
tʰaw²¹kɛ³³ 头家	thawkaaɛ ថៅកៃ	thaw⁴¹kɛ:²² เถ้าแก่	thâwkɛ̄ɛ	shopkeeper; tradesman; boss
tʰaw²¹kɛ³³niɔ⁵⁵ 头家娘		thaw⁴¹ke:³³nia⁴⁵³ เถ้าเกเนี้ย		wife of the boss
tʰaw²¹naŋ⁵⁵ 头人		thaw⁴¹naŋ⁴⁵³ เถ้านั้ง		a respected, elder Chinese
tia³⁵ 爹	tìiə ឪ	tia²² เตี่ย	tīə	father
tiɔ²¹tsiw³³ 潮州	tìiəciiw ឆាជីវ	tɛ:⁴¹tɕiw²⁴ แต้จิ๋ว	tɛ̄ɛchǐw	Teochew
tiŋ²¹naŋ⁵⁵ 唐人		tiŋ⁴¹naŋ⁴⁵³ ตึ้งนั้ง		Chinese person
tsɔ³²sũa³³ 座山		tɕaw⁴¹sua²⁴ เจ้าสัว		a rich man, esp. Chinese
tsek²² 叔	(ʔaa)cɔ̀k ចិក	tɕek⁵⁵ เจ๊ก	cēk	a Chinese man
tsiŋ³³teŋ³³ 亲丁		tɕi:n³³teŋ³³ จีนเต็ง		a trusted follower
tsuŋ²¹tsu⁵¹ 船主	cɔ̀ncuu ចិនជូ	tɕun⁴¹tɕu:⁴⁵³ จุ้นจู๋		the master of a ship
tua²¹puj⁵⁵ 肥大		tua⁴¹puj⁴⁵³ ตั้วปุ้ย		to be pot-bellied, fat
zi²¹pũa⁵⁵ 二盘		ji:⁴¹pua⁴⁵³ ยี่ปั๊ว		a second-level seller

The Teochew pronunciations and Chinese characters in Table 3.1 above come from Gyarunsut (1983). English translations have been adapted from Cooper (2020) and Pou & Jenner (1973). The Khmer and Lao pronunciations are from Pou & Jenner (1973) unless otherwise noted, but their ⟨q⟩ has been replaced by ⟨ʔ⟩. Thai script and pronunciations come from Gyarunsut (1983). Note that exact usages of these terms may vary between Teochew, Khmer, Thai, and the other languages, as is common with loanwords. English definitions have been combined for simplicity. See Cooper (2020) and Pou & Jenner (1973) for more nuanced usages.

3   Khmer Loanwords in Cambodian Teochew

The Chinese, with their long history in Cambodia, have influenced the Khmer vocabulary, yet the reverse is also true. This section explores Khmer loanwords in Cambodian Teochew, the language of the majority ethnic Chinese group in Cambodia.

3.1   *Methodology*

Data was gathered from nine speakers in Phnom Penh over four trips made in 2018–2020. Additionally, this chapter integrates data from two speakers consulted in McFarland (2017) who immigrated to the United States in the 1980s. Seven consultants were found through the researcher's familial connections in Cambodia and four were recruited via direct message requests targeted at Phnom Penh residents who were participants of Teochew-related Facebook pages. Informal interviews were conducted with each speaker, and any present family members, where they were asked about their educational and language backgrounds and speaking practices and ideologies. Three generations were identified based on the collected background information and historical events in Cambodia. Generation 1 (G1) consists of speakers over 65 who were born in Cambodia and grew up before the wartime and Vietnamese occupation (1970–89). Their parents were likely also born in Cambodia or moved there at a very young age. Five speakers were consulted in this category, four females and one male. Generation 2 (G2) is made up of those speakers' children, aged 35–65. These people were children during the wartime. Three speakers, one female and two males, were consulted in this category. Finally, Generation 3 (G3) represents the children of G2 who are under 35. These individuals often attended the Chinese schools that reopened in the 1990s. Three female speakers were consulted in this category. All consultants speak Teochew and Khmer, while five also speak English (2G2, 3G3,), and seven (2G1, 2G2, 3G3) speak some amount of Mandarin. Interviews were conducted in English with

G3 speakers and the two G2 speakers from McFarland (2017), and in a mixture of Khmer and Teochew with the G1 speakers and the other G2 speaker via the assistance of English-speaking family members.

A Zoom Q8 Handy video recorder and lapel and headset microphones were used with a 48 kHz sampling rate and 24-bit depth. Consultants were shown pictures of items and actions and asked to produce an appropriate word in Teochew. Prompts consisted of images of local foods and dishes, items from Swadesh's (1955) list of non-cultural vocabulary, and selected words that were listed as loanwords in Singapore Teochew by Goh (2017) and Low (2014). Naturalistic texts were also recorded by having consultants tell stories, describe scenes from picture books, and perform the family problem picture task from San Roque et al. (2012). Some additional data was collected from five of the speakers over Facebook Messenger. The assembled data constituted a preliminary vocabulary of Cambodian Teochew which was then compared to sources on other Teochew varieties including Indonesian (IT) (Peng 2012), Jieyang (JT) (Xu 2007), Singapore (ST) (Goh 2017; Li 1991; Low 2014; Yeo 2011), and Thai (TT) (Atchariyasucha 1982). References were also made to historical works on Teochew (HT) from foreign missionaries to China such as Ashmore (1884), Dean (1841), Duffus (1883), Fielde (1878), Fielde (1883), Giles (1877), Goddard (1888), and Lim (1886), as well as to a mobile Teochew dictionary application WhatTCSay (TCKnow LLC. 2015) and an online pronunciation dictionary Mogher.com. If a Cambodian Teochew word was found to be similar to the word form in another Teochew source, it was determined that it was likely not borrowed from Khmer. If a vocabulary item did not resemble the form from any of the other Teochew varieties, or if the word was not present at all in any of the other sources, its Khmer equivalent was found to determine if they were similar. Khmer words were found using online dictionaries and translation services. Comparable words were marked as likely loanwords. Given that Khmer has borrowed extensively from Chinese as shown in the previous section, a final check was made to Pou and Jenner (1973) to confirm the direction of the borrowing. For example, the Cambodian Teochew word $t\mathcal{c}^ha^{24}t\mathcal{c}^hau^{52}$ 吵吵 'to bother, pester, quarrel' resembles the Khmer word *chaachaw* with the same meaning. But according to Pou and Jenner, Khmer has borrowed this word from Chinese (1973: 47). Therefore, if a word was found in Pou and Jenner, it was decided that it was not a case of Cambodian Teochew borrowing from Khmer.

3.2   *Loanwords*
After going through the aforementioned process with the collected vocabulary items, it was determined that the items listed in Table 3.2 constitute words

TABLE 3.2   Khmer loanwords in Cambodian Teochew

Cambodian Teochew	English gloss	Khmer etymon[a]	Other Teochew sources
a¹¹mɔk⁵	amok (curry in banana leaf)	a:mɔk អាម៉ុក	
bɔk²lɔ¹¹hɔŋ⁵⁵	papaya salad	ɓɔklhɔŋ បុកល្ហុង	
bw³³	avocado	pʰlaɛɓɚ ផ្លែប័រ	ŋak³³laɪ⁵⁵ 鳄梨, gu¹¹ni¹¹kuɛ̃⁵³ 牛奶果 (TCKnow 2015)
bəŋ³³tɕʰau⁵²	Cambodian crepe, pancake	ɓaɲchaɛʊ បាញ់ឆែវ	
hɔu¹¹ka¹¹dou¹¹	give gift	kadoʊ កាដូ	HT loi⁵³-muɛʔ⁵, saŋ³¹³-loi⁵³, siɛ³³-saŋ³¹³ (Duffus 1883: 122), saŋ³¹³-loi⁵³-mueʔ⁵ 送禮物 (Fielde 1883: 465), JT li⁵³⁻³⁵mueʔ⁵ (Xu 2007: 180)
ka¹¹lem³³	ice cream	ka:reɛm ការ៉េម	suŋ³³kɔ³³ 霜膏, sɔʔ⁵kɔ³³ 雪糕 (TCKnow 2015), ST ai³³sə³³kikˡlim⁵⁵ 爱士吉林 (Goh 2017: 205)
ka³³laŋ³³	kralan (bamboo sticky rice)	krʊla:n ក្រឡាន	
lɔ¹¹hɔŋ⁵⁵	papaya	lhɔŋ ល្ហុង	HT bak⁵kue³³ 木瓜 (Fielde 1883: 277), IT pakue, nikue (Peng 2012: 33), ST bak⁵kue³³ 木瓜, ni³³kue³³ 奶瓜 (Goh 2017: 134)
lɔk²laʔ⁵, tɕʰa⁵²⁻²⁴lɔk²laʔ⁵	lok lak (stir-fried beef cubes)	lɔklaʔ ឡុកឡាក់	
maŋ¹¹kʰuʔ²	mangosteen (*Garcinia mangostana*)	mʊəŋkʰut មង្ឃុត	ST maŋ⁵⁵hek¹ 芒黑 (Goh 2017: 209)
pɔ³³la³³hɔk⁵	prahok (fermented fish paste)	prʊhɔk ប្រហុក	

a   Filippi and Hiep (2016) and Sak-Humphry (2015) were referenced for Khmer pronunciation.

TABLE 3.2 Khmer loanwords in Cambodian Teochew (*cont.*)

Cambodian Teochew	English gloss	Khmer etymon	Other Teochew sources
pɔ³³lɔi⁵²	jungle, forest	prei ព្រៃ	HT tʃʰi³¹³-nã⁵⁵-ta³¹³, tʃʰi³¹³-pʰɛ³¹³, tʃʰiu¹¹-bak² im³³-uɛ³¹³ (Duffus 1883: 158), tʃʰiu¹¹-lim⁵⁵, tʃʰiu¹¹-nã⁵⁵ (Duffus 1883: 114), kuaŋ ia 曠野 (Dean 1841: 18), tsʰiu¹¹lim⁵⁵ 樹林 (Fielde 1883: 354), ST tsʰɪu¹¹lim⁵⁵ (Yeo 2011: 12), TT lim⁵⁵ 林 (Atchariyasucha 1982: 121)
pi³³	from (time or place)	pi: ពី	HT taŋ¹¹ 從[b] (Fielde 1878: 172), tʰaŋ³¹³ 亘 (Fielde 1883: 576), tsʰoŋ⁵⁵, iu⁵⁵ (Ashmore 1884: 55; Duffus 1883: 117), tsʰoŋ⁵⁵ 從 (Goddard 1888: 24; Lim 1886: 5), JT tsʰoŋ⁵⁵ (Xu 2007: 31), ST taŋ¹¹ (Low 2014: 54),[c] tʰaŋ²¹³ 亘 (Goh 2017: 173)
piŋ³³pɔŋ³³	balloon	pe:ŋpaɔŋ ប៉េងប៉ោង	HT pʰu⁵⁵-hun⁵⁵-kiu⁵⁵, tʰiɛn³³-tsun⁵⁵, puɛ³³-tʃʰia³³, kʰi³¹³-kiu⁵⁵ (Duffus 1883: 17), ST be³³loŋ⁵⁵ 碼隆 (Goh 2017: 206)
sau³³mau³³	rambutan (*Nephelium lappaceum*)	sa:vma:v សាវម៉ាវ	ST aŋ⁵⁵⁻¹²mɔ⁵⁵⁻¹²taŋ³³ 红毛丹 (Li 1991: 59), aŋ⁵⁵-mo⁵⁵taŋ³³ 红毛丹 (Goh 2017: 205)

b Fielde (1878: 8) noted that this character is not of direct equivalence but is "of similar signification".

c Low (2014) did not include citation tones in her pitch superscripts. Since the one appearance of *taŋ¹¹* 'from' in her paper was in a sandhi environment, according to her tone sandhi rules, the citation form could be either *taŋ⁵⁵*, *taŋ³⁵*, or *taŋ¹¹*.

that Cambodian Teochew has borrowed from Khmer. The table shows the loanwords in Cambodian Teochew, their English meaning, the Khmer pronunciation, and the equivalents in other Teochew varieties if found in the literature. Chinese characters are provided next to the IPA if they were present in the original source.

3.3   *Integration of Loanwords*

Loanwords are adapted phonologically and morphologically to conform to the constraints of the borrowing language (Kang 2011). In terms of phonological adaptation, the words borrowed from Khmer adhere to the phonotactics of Cambodian Teochew. For example, while Khmer has both /l/ and /r/ in its phonological inventory (Huffman 1967), Cambodian Teochew has only /l/ (Chen 2009; McFarland 2017; Pan 2000). Therefore, Khmer /r/ is pronounced as /l/, as seen in the words for 'jungle', 'prahok', and 'ice cream'. Additionally, while Khmer allows consonant clusters at syllable onset as seen in *prei* 'jungle', *prɒhɔk* 'prahok', *lhɔŋ* 'papaya', and *krɒlaːn* 'kralan', these are not allowed in Teochew. To resolve this, the loanword could delete one of the two consonants or insert an epenthetic vowel (Miao 2005:105). In the first three instances, we see epenthetic vowel insertion between the two consonants to get $pɔ^{33}lɔi^{52}$, $pɔ^{33}la^{33}hɔk^{5}$, and $lɔ^{11}hɔŋ^{55}$. For 'kralan', the /r/ is deleted to get $ka^{33}laŋ^{33}$. Cambodian Teochew also restricts consonants in syllable coda and /n/ is not allowed. We can see that the final letter of Khmer *krɒlaːn* 'kralan' is pronounced /ŋ/ in the Teochew word $ka^{33}laŋ^{33}$. Finally, while Khmer is non-tonal, Cambodian Teochew is a tonal language, so each borrowed word gets assigned a tone, as reflected by the pitch values given in the numerical superscripts.[7] Teochew also has extensive tone sandhi, a phenomenon common in Sinitic languages, where a word in a phrase or sentence will surface with a different tone than when the word appears in isolation. The general rule is that the right-most syllable in a word or utterance retains its citation tone while the other syllables are subject to tone sandhi.[8] For example, hue^{52} 'fire' becomes $hue^{52-24}suã^{33}$ 'volcano'. Per convention, a word's dictionary/citation tone appears first in the superscript, followed by a hyphen, and then the word's surface tone after undergoing tone sandhi. There is evidence that the Khmer loanwords in Cambodian Teochew undergo tone sandhi as the phrases in (1) show. Similarly, they provide an environment

---

7   Variation may exist amongst speakers in the word's tone. Even still, borrowed words appear to have been given tones and follow rules of tone sandhi as outlined below.

8   Not all words surface in different tones. Words with mid-level (pitch value: 33) and low-level (pitch value: 11) tones remain unchanged in sandhi domains. See McFarland (forthcoming) for a full overview of the tone and tone sandhi patterns in Cambodian Teochew.

for tone sandhi such that the preceding words surface in their sandhi tones, as shown in (2).

1) a. lɔ¹¹hɔŋ⁵⁵⁻¹¹   tɕui⁵²
      ល្ហុង   水
      papaya   water
      'papaya juice'

   b. a¹¹mɔk⁵⁻²   hɯ⁵⁵
      អាម៉ុក   鱼
      amok   fish
      'amok fish'

2) a. tɕʰa⁵²⁻²⁴   lɔk²laʔ⁵
      炒   ឡុកឡាក់
      stir-fry   lok lak
      'stir-fried beef cubes'

   b. i³³   nim²⁴⁻¹¹   piŋ³³pɔŋ³³
      伊   □   ប៉េងប៉ោង
      3SG   hold   balloon
       'She holds the balloon.'

In terms of morphological adaptation, Cambodian Teochew is like other Sinitic languages in having limited morphology, so integration is relatively straightforward. Borrowed words do not need to be inflected for plurality or gender. In Teochew (and also in Khmer), nouns are counted and referred to with a classifier. The word order is number, classifier, noun. The classifier used is dependent on the type of noun. For example, *tɕiaʔ²* 隻 is used with animals and *bue⁵²* 尾 with fish. My consultants provided evidence that Khmer loanwords in Cambodian Teochew can be used with a specific classifier like *liap⁵* 粒, used for round things, as in (3), or with the generic classifier *kai⁵⁵* 个, which can also be used with most things, as shown in (4).

3) a. sã³³   liap⁵⁻²   lɔ¹¹hɔŋ⁵⁵
      三   粒   ល្ហុង
      three   CLF   papaya
      'three papayas'

b. sã³³   liap⁵⁻²   maŋ¹¹kʰuʔ²
三   粒   មង្ឃុត
three   CLF   mangosteen
'three mangosteens'

c. ŋɔ²⁴⁻¹¹   liap⁵⁻²   sau³³mau³³
两   粒   សាវម៉ាវ
two   CLF   rambutan
'two rambutans'

4)   a. sã³³   kai⁵⁵⁻¹¹   lɔ¹¹hɔŋ⁵⁵
三   个   ល្ហុង
three   CLF   papaya
'three papayas'

b. sã³³   kai⁵⁵⁻¹¹   maŋ¹¹kʰuʔ²
三   个   មង្ឃុត
three   CLF   mangosteen
'three mangosteens'

c. ŋɔ²⁴⁻¹¹   kai⁵⁵⁻¹¹   sau³³mau³³
两   个   សាវម៉ាវ
two   CLF   rambutan
'two rambutans'

d. si¹¹⁻⁵³   kai⁵⁵⁻¹¹   piŋ³³pɔŋ³³
四   个   ប៉ុងប៉ោង
four   CLF   balloon
'four balloons'

Khmer also has limited morphology, though one prefix used is *pʰlaɛ* meaning 'fruit'. The word *pʰlaɛɓɚ* 'avocado' can be translated to 'fruit-butter'. In isolation, *ɓɚ* is a French borrowing in Khmer meaning 'butter'. The word *pʰlaɛ* can be used with fruits like papaya (*pʰlaɛlhɔŋ*), mangosteen (*pʰlaɛmʋəŋkʰut*), and rambutan (*pʰlaɛsaːʋmaːʋ*). For these words, *pʰlaɛ* is not obligatory as the words without said prefix do not have another meaning, unlike with 'avocado' (*pʰlaɛɓɚ*) and 'butter' (*ɓɚ*). Importantly, when Cambodian Teochew borrows the Khmer word, *pʰlaɛ* is never used. The loanword for avocado becomes solely

bɯ³³. With these mentioned stipulations, loanwords from Khmer integrate seamlessly into the Cambodian Teochew language.

3.4 *Speaker Variation*

It should be noted that not all of the eleven speakers interviewed for this chapter necessarily use all the words in Table 3.2. Table 3.3 shows a breakdown of the number of speakers in the gathered data using each word and the alternate words used, if any. 'No word' signifies that the speaker was shown a prompt for the item and they indicated they did not know the word in Teochew. 'Unknown' signifies that it is not known if the speaker uses the word in their Teochew. This is due to an expanding list of prompts in later speaker consultations and the researcher's inability to re-interview speakers from early consultations. Though this study did not have enough participants to present anything conclusive on gender and generational variations, any points of interest related to speaker age will be covered in section 4. At least three speakers needed to use the loanword for it to be included in Table 3.2. 'Jungle', 'mangosteen', 'rambutan', 'papaya', 'papaya salad', 'prahok', 'pancake/crepe', 'ice cream', 'balloon', and 'from' are also attested by the Cambodian Teochew diaspora in the Facebook group *Gaginang*.

TABLE 3.3   Breakdown of the count of speakers using each word

English gloss	Word used	Count, generation, gender
jungle, forest	7 pɔ³³lɔi⁵²	2G1F, G1M, G2F, 2G2M, G3F
	1 pɔ³³lɔi⁵² + tɕʰiu¹¹lim⁵⁵	G1F
	1 tɕʰiu¹¹lim⁵⁵	G1F
	2 no word	2G3F
mangosteen	10 maŋ¹¹kʰuʔ²	4G1F, G2F, 2G2M, 3G3F
	1 unknown	G1M
rambutan	10 sau³³mau³³	4G1F, G2F, 2G2M, 3G3F
	1 unknown	G1M
avocado	4 bɯ³³	3G1F, G3F
	1 gu⁵⁵⁻¹¹ni¹¹kue⁵²	G3F
	1 gu⁵⁵⁻¹¹iu⁵⁵⁻¹¹kue⁵²	G2M
	1 ɑːvəkʰɑːdoʊ	G3F
	3 no word	G1F, G2F, G2M
	1 unknown	G1M
papaya	10 lɔ¹¹hɔŋ⁵⁵	4G1F, G2F, 2G2M, 3G3F
	1 unknown	G1M

TABLE 3.3   Breakdown of the count of speakers using each word (*cont.*)

English gloss	Word used	Count, generation, gender
papaya salad	11 bɔk^{2}lɔ^{11}hɔŋ^{55}	all
prahok	11 pɔ^{33}la^{33}hɔk^5	all
Cambodian crepe	11 bəŋ^{33}tɕʰau^{52}	all
ice cream	7 ka^{11}lem^{33}	3G1F, G2F, G2M, 2G3F
	1 siɔʔ$^{2–5}$kɔ^{33}	G3F
	1 səŋ^{33}kɔ^{33}	G1F
	2 unknown	G1M, G2M
kralan	3 ka^{33}laŋ^{33}	G1F, 2G3F
	2 tek$^{2–5}$kɔŋ^{33}kue^{52}	G1F, G2M
	2 no word	G2F, G3F
	4 unknown	2G1F, G1M, G2M
lok lak	8 lɔk^{2}laʔ^5	2G1F, G1M, G2F, 2G2M, G3F
	3 tɕʰa$^{52–24}$lɔk^{2}laʔ^5	2G1F, G3F
	1 unknown	G3F
amok	9 a^{11}mɔk^5	4G1F, G2F, 2G2M, 2G3F
	2 unknown	G1M, G3F
to give a gift	3 hou^{11}ka^{11}dou^{11}	3G3F
	2 saŋ^{53}lɔi^{52}	G2F, G2M
to gift something	2 saŋ^{53}mueʔ^{2}kiã^{35}	2G1F
to gift a book	1 saŋ^{53}pou^{24}	G1F
to give	3 hou^{11}	G1F, G1M, G2M
from	3 pi^{33}	3G3F
	6 unknown	4G1F, G1M, G2F, 2G2M
balloon	6 piŋ^{33}pɔŋ^{33}	G1F, G2F, G2M, 3G3F
	5 unknown	3G1F, G1M, G2M

4   Discussion

Selinker (1992) categorized lexical changes in language contact situations. Terms can be classified as expansive (a new word fills a gap in the lexicon), additive (new and old terms are both used), replacive (the former word disappears), loan shift (an old word's meaning changes to fill a lexical gap), loan translation (new words or phrases are translated literally), or loan blend (the term combines words or parts of words from multiple languages)

(Selinker 1992: 46). The data in Tables 3.2 and 3.3 provide evidence for expansive, additive, replacive, and loan blend loanwords in Cambodian Teochew.

Expansive vocabulary would be terms for local dishes like 'papaya salad', 'Cambodian crepe', 'prahok', 'kralan', 'amok', and 'lok lak' that likely did not exist in the language of the historic Teochew settlers in Cambodia. The Khmer word may have been adopted out of necessity and/or convenience. 'Papaya salad', 'Cambodian crepe', 'prahok', 'amok', and 'lok lak' were strongly attested in the data (by nine or more speakers), and no other words were provided as alternatives to the Khmer loanword. The evidence was not as strong for 'kralan', a dish made of sticky rice inside bamboo. Only three speakers (G1, 2G3) said the Khmer loanword $ka^{33}laŋ^{33}$. Two speakers (G2, G3) did not know a Teochew word for it and two (G1, G2), said $tek^{2-5}kɔŋ^{33}kue^{52}$ 竹管粿 which translates to 'bamboo cake'. The use of this word suggests that historic settlers in Cambodia came across 'kralan' and created a new word for it using terms already present in their Teochew like tek^2 竹 'bamboo' and kue^{52} 粿 'cake', rather than adopting the Khmer term.[9]

Other potentially expansive vocabulary items include the tropical fruits 'rambutan' and 'mangosteen'. The English word 'rambutan' is itself a loanword from Malay *rambutan*. This fruit is often red in colour with hair-like protuberances. Singapore Teochew's $aŋ^{55}mo^{55}taŋ^{33}$ (Goh 2017: 205) (and Mandarin's $hoŋ^{35}mao^{35}tan^{55}$ 红毛丹) was created with phono-semantic matching (see Ding, this volume); $aŋ^{55}mo^{55}$ 'red hair' describes the appearance of the fruit and the $taŋ^{33}$ is a phonetic matching of the final syllable of the Malay word (where /n/ is pronounced /ŋ/ due to Teochew phonotactics). Similarly, 'mangosteen' comes from Malay *maŋgis*. Singapore Teochew has again used phono-semantic matching with the Malay word in $maŋ^{55}hek^1$ 芒黑 (Goh 2017: 209), where $maŋ^{55}$ matches the source word's pronunciation and hek^1 'black, dark' describes the fruit's outer colour. It is not known if the early Teochew settlers to Cambodia had words for 'rambutan' or 'mangosteen'. If they did and adopted the Khmer loanwords $sau^{33}mau^{33}$ and $maŋ^{11}k^hu\textipa{P}^2$ (each used by ten speakers in this study), these would be examples of replacive borrowing. If terms did not exist, they would be additional cases of expansive borrowing.

Food items such as 'avocado' and 'ice cream' are also likely expansive loanwords, though these are not native to Khmer cuisine. No word was found for 'avocado' in the historical sources. It is likely a new food item introduced after Teochew speakers had arrived in Cambodia. The Khmer loanword bu^{33}, attested in four speakers (3G1, G3), was likely adopted out of efficiency. Alternate words include $gu^{55-11}ni^{11}kue^{52}$ 牛奶果 ('milk fruit') used by

---

9   Thus $tek^{2-5}kɔŋ^{33}kue^{52}$ is also an example of an expansive vocabulary item.

one G3 speaker, and *gu$^{55-11}$iu$^{55-11}$kue^{52}* 牛油果 ('butter fruit') used by one G2 speaker. In these cases, existing Teochew words were put together to create the new word 'avocado'.[10] The Khmer word 'ice cream' *ka:reɛm* is borrowed from French *crème*. Since the Cambodian Teochew form *ka^{11}lem^{33}* so closely resembles the Khmer pronunciation, it was determined that the Khmer form likely influenced its pronunciation in the language of interest. Seven speakers in this study used this borrowed form. Alternatively, one speaker said *siɔʔ$^{2-5}$kɔ^{33}* 雪糕 which translates literally to 'snow cake'. In contrast, Singapore Teochew uses the loanword *ai^{33}sə^{33}kik^{l}lim^{55}* 爱士吉林 from English 'ice cream' (Goh 2017: 205).

The data provided evidence of additive loanwords in Cambodian Teochew, e.g. the word for 'jungle, forest'. Dean (1841) is the only historical source with a word glossed as 'jungle', while Duffus (1883) and Fielde (1883) have words listed under 'forest'. Either way, Cambodian Teochew speakers seemingly chose to adopt the Khmer word for 'jungle, forest' in *pɔ^{33}lɔi^{52}*, used by eight speakers in this study. Per Haspelmath (2009: 49), "lexical meanings do not have to fit into predefined slots". Potentially the historic Teochew words for 'forest' were not sufficient to describe the wooded areas in Cambodia. Consultants were shown several pictures at different times of various types of wooded areas in an attempt to disconcert a difference between 'forest' and 'jungle'. Evidence points to some usage of historic Teochew terms, in addition to the adoption of a Khmer loanword. As Table 3.3 shows, one speaker (G1) used both *pɔ^{33}lɔi^{52}* and the historic *tɕ^{h}iu^{11}lim^{55}*, while one speaker (G1) used solely the latter term. Two G3 speakers never produced a Teochew word for 'jungle' or 'forest', which hints at its rarer usage in Cambodia.[11] Khmer *prei* ព្រៃ 'forest' has metaphorical importance in Cambodian culture (see Edwards 2008 and Lim 2011), so it is perhaps for this reason that the Cambodian Teochew have adopted the Khmer word into their language. *Lim^{55}* 林 'forest' is the second most common surname in the Chaoshan region, held by over 1 million people (The Teochew Store). This is perhaps why the historic Teochew word has not been eschewed altogether.

---

10   *gu$^{55-11}$iu$^{55-11}$kue^{52}* 牛油果 'butter fruit' could also be a loan translation from Khmer *p^{h}laɛɓ♂* 'fruit-butter'.

11   Singapore Teochew speakers also had difficulty producing a word for 'forest' in Ho's (2009) study on vocabulary retention in younger speakers. Ho (2009) found that words related to nature including 'forest', 'lake', and 'fog' were commonly missing in the vocabulary of the speakers in the study. My data provides evidence that this may also be the case for some speakers in Cambodian Teochew. In addition to 'forest' which was explained above, one G3 speaker used a loanword *bəŋ^{33}* 'lake' from Khmer *ɓɔŋ* បឹង (versus historic Teochew *ɔu^{55}* 湖 used by three speakers, 2G3F, G1F). Meanwhile, the two other G3 speakers used a loanword *ap^{5}* for 'fog' from Khmer *ʔap* អ័ព្ (versus historic Teochew terms *mɔŋ^{55}* 蒙, *mɔu^{55}* 雺, and *bu^{11}* 霧 used by G1F, G1F, and G3F respectively).

Cambodian Teochew *pi³³* 'from', which comes from Khmer *pi:* ពី, is the one example of the borrowing of a function word.[12] Generally, we can see *taŋ¹¹* 'from' used in time expressions in Singapore Teochew (5a) and historical Teochew (5b), and in locative expressions in historical Teochew in (6).

5) a. a³³-Zach   taŋ¹¹   sɔi²¹³   tsu¹¹   tsiŋ³³   tsĩa¹¹gau¹¹   puaʔ²   lai⁵⁵
        NVOC-PN   from   young   SEQ   very   good.at   fall   come
        puaʔ²   khɨ¹¹
        fall   go
        'From a young age, Zach was very accident-prone.' (Low 2014: 54)

    b. taŋ¹¹   tsʰuʔ²   si³¹³   kau³¹³   tã³³
        from   birth   arrive   now
        'from his birth till now' (Fielde 1878: 172)

6) taŋ¹¹   tsi⁵³   siã⁵⁵   kau³¹³   nin⁵³   tse¹¹   lai²⁵   u²⁵
    from   this   city   arrive   2PL   village   there is
    dzieʔ⁵⁵   tsoi²⁵   li⁵³   lo²⁵
    how   much   distance
    'How far is it from this city to your village?' (Fielde 1878: 240)

In Khmer, *pi:* 'from' is also used with locative (7) and time (8) expressions.

7) mɔɔk   pi: srok   ʔɑŋkleeh
    មក   ពី   ស្រុក   អង់គ្លេស
    come   from   country   England
    '(She) comes from England.' (Sak-Humphry 2015: 5)

8) cap   pi: moaŋ   pram   kɑnlah   rɔhoot dɑl   moaŋ   dɑp   yup
    ចាប់   ពី   ម៉ោង   ប្រាំ   កន្លះ   រហូតដល់   ម៉ោង   ដប់   យប់
    starting   from   hour   five   half   until   hour   ten   night
    'starting from 5:30PM until 10:00PM' (Sak-Humphry 2015: 50)

There is evidence that loanword *pi³³* 'from' in Cambodian Teochew is similarly used with both locative (9) and time expressions (10).

---

12   Function words are generally less borrowable than content words (Tadmor 2009). There is evidence of other function words borrowed in Teochew varieties in the region including *itu* 'this' and relativizer *yang* in Indonesian Teochew (Peng 2012) and *ta³³pɨʔ⁵* 'but' in Singapore Teochew (Yeo 2012), all borrowed from Malay.

9) i³³   pe⁵⁵⁻¹¹   tɕʰuk²⁻⁵   pi³³   paŋ⁵⁵
    伊   爬   出   ㆠ   房
    3SG   crawl   exit   from   room
    'He crawled out from his room.'

10) pi³³   ik²⁻⁵   tiam⁵²   kau¹¹⁻⁵³   ŋou²⁴⁻¹¹   tiam⁵²
    ㆠ   一   点   到   五   点
    from   one   o'clock   arrive   five   o'clock
    'from 1 o'clock to 5 o'clock'

Evidence shows that Cambodian Teochew has replacive borrowings in 'papaya' and 'balloon'. The word *bak⁵kue³³* 木瓜 'papaya' was found in Fielde (1883: 277) and in Indonesian and Singapore Teochew varieties. No evidence was found for the existence of *bak⁵kue³³* in Cambodian Teochew. Conversely, *lɔ¹¹hɔŋ⁵⁵* 'papaya' was used by ten speakers in this study. The author suggests that the use of the Khmer-origin word comes as an extension of the loanword for the popular 'papaya salad' dish, *bɔk²lɔ¹¹hɔŋ⁵⁵*. Since the Cambodian Teochew were already using *lɔ¹¹hɔŋ⁵⁵* for this dish's name, it would be more efficient to refer to the fruit 'papaya' using the same word. Duffus (1883: 17) presented several words for 'balloon' including *pʰu⁵⁵-hun⁵⁵-kiu⁵⁵*, *tʰiɛn³³-tsun⁵⁵*, and *kʰi³¹³-kiu⁵⁵*. Each of these words has the morpheme *kiu⁵⁵* 球 'ball'. Per Weinreich (1979), relatively infrequent vocabulary items are more likely candidates for replacement. This may be the case with 'balloon' and so the Khmer word was adopted out of convenience. The Khmer loanword was attested by six speakers (G1, 2G2, 3G3) and no evidence was found for the use of historic Teochew words. Singapore Teochew has also adopted a loanword *be³³lɔŋ⁵⁵* 码隆 from English (Goh 2017: 206).

Finally, Cambodian Teochew has one loan blend in 'to give a gift'. This is a hybrid Teochew-Khmer phrase in which the verb 'give' is pronounced as Teochew *hɔu¹¹* 互, while the noun 'gift' *ka¹¹dɔu¹¹* comes from Khmer *kadoʊ* កាដូ 'gift'; *hɔu¹¹ka¹¹dɔu¹¹* is specifically used to mean to give someone a gift or present. No evidence was found of *ka¹¹dɔu¹¹* being used in other contexts and it always appeared with *hɔu¹¹*. The three G3 speakers in this study used this phrase. The picture shown to consultants was an extended hand with a box with a bow. Speaker responses to this cue varied. Two speakers said solely *hɔu¹¹* when shown the prompt; *hɔu¹¹* 'to give' is used in Cambodian Teochew in phrases such as 'to give money' or 'to give plants'. Five speakers used a phrase with *saŋ⁵³* 送, which means 'to gift'.[13] Two speakers said *saŋ⁵³lɔi⁵²* 送礼 'to give

---

13   Citation tone for this word in Cambodian Teochew is unconfirmed.

a gift', two *saŋ⁵³mueʔ²kiã²⁴* 送物件 'to give something' and one *saŋ⁵³pɔu²⁴* 送簿 'to gift a book'. Given that it was the younger speakers who used the loan blend, it could be a relatively new word adaptation in Cambodian Teochew. More evidence is needed to verify this claim.

5 Concluding Remarks

This section discusses some limitations to the current study and summarize its findings and implications, while also providing directions for future research. The fifteen words found in this chapter and presented in Table 3.2 do not constitute an exhaustive list of Cambodian Teochew's Khmer loanwords. More examples might surface with additional fieldwork and data analysis. Because the methodology relied on elicitation using wordlists, and descriptions of fairly simple pictorial stories, the collected vocabulary items are limited. There are likely many more words that are beyond the categories this current study has covered.[14]

Furthermore, given the relatively few speakers consulted for this project, and other reasons outlined below, it is impossible to generalize about Cambodian Teochew as a whole and this chapter attempted to only present findings based on the collected data. Variation existed even amongst the eleven consultants, as shown in Table 3.3, so more variation would likely be found in a larger sample size. Because seven speakers in this project were related, though they were living in four different households, there was the potential that the data would show the Teochew spoken by only one family. Attempts made to diversify participants by recruiting from online spaces succeeded in bringing in an additional four speakers, all from separate families. Yet the online recruiting process itself likely influenced the type of speakers integrated in this project. Online spaces included Facebook group *Gaginang* where interactions are primarily conducted in English, and the Cambodian Teochew Association's Facebook page which is primarily conducted in Mandarin. Someone not speaking one of those languages, or someone not on Facebook, may not have been a participant in one of those spaces, thus making it challenging for them to be included in this study. However, one speaker found through this method was not a participant herself, and instead was introduced to me through her grandchildren. As for the three other speakers recruited on Facebook, it is

---

14 For example, Teo (1993) found the Peranakan Chinese in Malaysia to use loanwords in areas such as diseases and illnesses, adjectives, and more rarely used nouns and verbs such as 'affair', 'curfew', 'sue', and 'assassinate', among others.

possible that their Teochew had degraded due to their use of other languages and thus their usage of loanwords from Khmer may be different from the general population. Additionally, all speakers interviewed were from Phnom Penh so there is no data on loanword usage by Cambodian Teochews living outside the capital city. This study could be expanded upon by working with more speakers with a broader set of picture prompts and stories and re-interviewing speakers from early consultations in order to check their loanword usages.

Challenges in the interview process have also affected data collection. First, there is the case of the observer's paradox in that my presence in the interviews would have affected the speakers' responses. For example, it may have been the case that speakers knew both a Khmer loanword and a historic Teochew word for a prompt, and even if they use the loanword more often in daily speech, they may have said the Teochew word in the interview because they knew my research was on that language. Furthermore, six interviews were conducted with the use of family members as interpreters which did not give me the same control as the interviews conducted solely by the researcher in English. Picture prompts were used in all interviews in order to reduce the effects of second language interference. Yet sometimes the interpreters would still say the target word in Khmer or in Teochew. This may have affected the speakers' responses. The consultant may have been more likely to say a Khmer loanword due to a Khmer prime, and their word choice may have similarly been affected if given a Teochew prime.

Fourteen of the fifteen words in Table 3.2 are content words versus function words and these fourteen are all nouns rather than verbs. This aligns with the findings from the Loanword Typology Project (Tadmor 2009). Furthermore, the loanwords suggest that lexical borrowing in Cambodian Teochew is primarily cultural.[15] However, Cambodian Teochew seems resistant to lexical changes overall. In comparison, Goh (2017) identifies over 200 foreign language loanwords in Singapore. The sociolinguistic situation in Singapore is very different from Cambodia, with that Teochew variety coming into contact with English, Hokkien, Malay, Mandarin, Tamil, and others, all of which are source languages for words on Goh's list. A likely contributing factor to the relatively few Khmer loanwords in Cambodian Teochew is the fact that many Khmer terms already resembled Teochew words, as explored in section 2. Thus staple foods like *mì* 面 'noodles' and *kûe tíaw* 粿条 'noodle soup' sound very similar to Khmer's *mii* មី and *kuj tiiəw* គុយទាវ respectively, since they were historically borrowed from Teochew (Pou & Jenner 1973). Even still, Thai similarly has

---

15   As opposed to core vocabulary (cf. Myers-Scotton 1993; Haspelmath 2009).

borrowed extensively from Teochew or other Sinitic varieties and Lin (2006) found over 60 Thai loanwords in the Teochew spoken in Thailand. Thus the findings in this chapter speak volumes to the Cambodian Teochews' determination to maintain their language and resist lexical changes, despite constant linguistic pressure from the dominant Khmer language.

References

Ashmore, William, *Primary Lessons in Swatow Grammar* (English Presbyterian Mission Press, 1884).

Atchariyasucha, Wichet, *A Phonological study of Swatow of Chinese as spoken in Bangkok with comparisons to Thai* (MA thesis, Mahidol University, Mahidol University Library and Knowledge Center, 1982). http://mulinet11.li.mahidol.ac.th/e-thesis/scan/14050.pdf.

Aw, Tash, 'Coming out of the shadows: what it means to be French and Chinese', *The Guardian* (2019, November 26). https://www.theguardian.com/world/2019/nov/26/what-you-hear-about-chinese-people-in-france-feeling-scared-its-true.

Bao, Zhiming, 'Tonal contour and register harmony in Chaozhou', *Linguistic Inquiry*, 30/3 (1999): 485–93.

Birnie-Smith, Jessica Rae, 'Ethnic identity and language choice across online forums', *International Journal of Multilingualism*, 13/2 (2016): 165–83. https://doi.org/10.1080/14790718.2015.1078806.

Bourgerie, Dana Scott, 'Education in the Cambodian Chinese diaspora', in O.E. Kagan, M.M. Carreira and C.H. Chik (eds.), *The Routledge Handbook of Heritage Language Education: From Innovation to Program Building* (London: Routledge, 2017): 161–78. https://doi.org/10.4324/9781315727974-13.

Chan, Sambath, *The Chinese minority in Cambodia: Identity construction and contestation* (MA thesis, Concordia University, 2005). https://spectrum.library.concordia.ca/8439/.

Chen, Shihlun Allen, *Socializing Chineseness: Cambodia's ethnic Chinese communities as a method* (PhD dissertation, University of Hawai'i at Manoa, publication number 10002238, 2015).

Chen, Xiaojin, '陈晓锦, 越南、柬埔寨、老挝三国潮州话训读现象比较 Comparative study on xundu phenomenon of Chaozhou dialect in Vietnam, Cambodia and Laos', 广东技术师范学院学报, 4 (2009): 79–84.

Chin, James K., 'Ethnicized networks and local embeddedness: The new Chinese migrant community in Cambodia', in M. Zhou (ed.), *Contemporary Chinese diasporas* (Berlin: Springer, 2017): 187–206. https://doi.org/10.1007/978-981-10-5595-9_9.

Chokkajitsumpun, Pranee, *Chinese literacy maintenance and shift in Bangkok: Individual and family cases* (unpublished PhD dissertation, University of Hawaiʻi at Mānoa, 1998).

Clayton, Thomas, *Language choice in a nation under transition: English language spread in Cambodia* (Berlin: Springer, 2006). https://doi.org/10.1007/0-387-31194-7.

Cole, Peter, Hermon, Gabriella and Lee, Cher Leng, 'Grammatical and discourse conditions on long distance reflexives in two Chinese dialects', in P. Cole, G. Hermon and C.-T.J. Huang (eds.), *Long distance reflexives* (Leiden: Brill, 2000): 1–46. https://doi.org/10.1163/9781849508742_002.

Cole, Peter and Lee, Cher Leng, 'Locality constraints on yes/no questions in Singapore Teochew', *Journal of East Asian Linguistics*, 6/2 (1997): 189–211. https://doi.org/10.1023/A:1008237705178.

Cooper, Doug, *Notes on Modern Chinese loanwords in Thai*, 2020. Retrieved June 1 from http://sealang.net/thai/chinese/modern.htm.

Dean, William, *First lessons in the Tie-chiw dialect* (Bangkok: Siam, 1841).

Dorais, Louis-Jacques, 'Refugee adaptation and community structure: The Indochinese in Quebec City, Canada', *International Migration Review*, 25/3 (1991): 551–73. https://doi.org/10.1177/019791839102500305.

Duffus, William, *English-Chinese vocabulary of the vernacular or spoken language of Swatow* (Swatow: English Presbyterian Mission Press, 1883).

Edwards, Penny, 'Between a song and a *prei*: Tracking Cambodian history and cosmology through the forest', in A.R. Hansen and J. Ledgerwood (eds.), *At the edge of the forest: Essays on Cambodia, history, and narrative in honor of David Chandler* (Ithaca: Cornell University Press, 2008): 137–62.

Edwards, Penny, 'Ethnic Chinese in Cambodia', in *Ethnic groups in Cambodia* (Phnom Penh: Center for Advanced Study, 2009): 174–234.

Edwards, Penny, 'Addendum to the introduction: New directions', *Cross-Currents: East Asian History and Culture Review*, 1/2 (2012): 274–78. https://doi.org/10.1353/ach.2012.0013.

Egerod, Søren, 'Swatow loan words in Siamese', *Acta Orientalia*, 23 (1958): 137–56.

Eiampailin, Janida, *The phonological interference of Swatow in Standard Thai by Chinese speakers in Bangkok* (MA thesis, Mahidol University, 2004).

Fielde, Adele Marion, *First lessons in the Swatow dialect* (Swatow: Swatow Printing Office Company, 1878).

Fielde, Adele Marion, *A pronouncing and defining dictionary of the Swatow dialect: Arranged according to syllables and tones* (Swatow: American Presbyterian Mission Press, 1883).

Filippi, Jean-Michel, 'Cambodia: The swing of the pendulum', *Phnom Penh Post*, 26 November 2010. https://www.phnompenhpost.com/post-plus/cambodia-swing-pendulum.

Filippi, Jean-Michel and Hiep, Chan Vicheth, *Khmer pronouncing dictionary: Standard Khmer and Phnom Penh dialect* (KAM, UNESCO, 2016). https://www.publikam.com/en/pages/bookstore-3.html.

Fishman, Joshua A., *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages* (Bristol: Multilingual Matters, 1991).

Giles, Herbert Allen, *Handbook of the Swatow dialect: With a vocabulary* (published with the assistance of the Straits' Government, 1877).

Goddard, Josiah, *A Chinese and English vocabulary, in the Tie-chiu dialect* (2nd edn., Swatow: American Presbyterian Mission Press, 1888).

Goh, Eng Choon 吴英俊, 潮州口语 *Colloquial Teochew* (Singapore: NETUCC, 2017).

Goh, Rui Long 吴芮珑, 端华学校: 柬埔寨华文教育的案例 *Duan Hua (Toun Fa) Chinese school: A case of Chinese education in Cambodia* (final year project, Nanyang Technological University, 2011). Digital Repository of NTU. http://hdl.handle.net/10356/43682.

Gong, Gunhu, *A list of Swatow loanwords in Thai* (unpublished PhD dissertation, Shantou University, 2000).

Gyarunsut, Pranee ปราณี กาย อรุณ สิทธิ์, คำ ยืม ภาษา จีน ใน ภาษา ไทย ปัจจุบัน *Chinese loanwords in modern Thai* (MA thesis, Chulalongkorn University, 1983). Chulalongkong University Intellectual Repository. http://cuir.car.chula.ac.th/handle/123456789/37061.

Haspelmath, Martin, 'Lexical borrowing: Concepts and issues', in M. Haspelmath and U. Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook* (Berlin: De Gruyter Mouton, 2009): 35–54. https://doi.org/10.1515/9783110218442.35.

Ho, Pei Qin 何佩芩, 语言的消亡: 以新加坡潮州话为个案: 兼论基本词汇的有阶分布 *Language decline: a case study on Singapore Teochew dialect and a study of the rank theory in basic word list* (final year project, Nanyang Technological University, 2009). Digital Repository of NTU. https://hdl.handle.net/10356/14980.

Hong, Ying, *A phonetic study of Chaozhou Chinese* (PhD thesis, Hong Kong University of Science and Technology, 2013). HKUST SPD: The Institutional Repository. http://hdl.handle.net/1783.1/62308.

Huang, Yu 黄瑜 and Mo, Yuan-Yuan 莫源源, '柬埔寨语中汉借词的语音特征分析 Phonological features of Chinese loanwords in Khmer', 课程教育研究 *Course Education Research*, 21 (2017): 95–97.

Huffman, Franklin Eugene, 'An outline of Cambodian grammar' (PhD thesis, Cornell University, 1967; publication number 6717239).

Kang, Yoonjung, 'Loanword phonology', in M. van Oostendorp, C.J. Ewen, E.V. Hume and K. Rice (eds.), *The Blackwell companion to phonology* (Hoboken: John Wiley, 2011): 1–25. https://doi.org/10.1002/9781444335262.wbctp0095.

Khoo, Kiak Uei 邱克威, '马来西亚雪兰莪洲滨海潮州渔村澄海人方言音系比较分析 A comparative phonological study of Chenghai dialects of Teochew fishing villages

in Selangor Malaysia', 马来亚大学华人文学与文化学刊第五卷 *Journal of Chinese Literature and Culture*, 5/1 (2017a): 14–24. https://ejournal.um.edu.my/index.php/JCLC/article/view/17465.

Khoo, Kiak Uei 邱克威, '马来西亚汉语方言声调变异及其社会因素的调查研究: 以霹雳州北部三个相邻渔村的普宁话为个案分析 Survey on the Malaysian Chinese dialect tonal variations and its social influences: Case study on Chaozhou (Puning) dialect of three neighbouring fishing villages in North Perak', 海外华文教育 *Overseas Chinese Education*, 2 (2017b): 258–67.

Khoo, Kiak Uei 邱克威, 馬來西亞巴冬潮汕話陰上調的方言融合與條件變讀 Dialect mixing and conditioned variants in *Yin Shang* tone of Chaoshan dialects in Parit Jawa fishing village of Malaysia', 臺灣語文研究, 13/2 (2018): 201–12. https://doi.org/10.6710/JTLL.201810_13(2).0002.

Kyne, Phelim, 'Cambodia's middle kingdom: Chinese schools, back from the brink', *Phnom Penh Post*, 25 June 1999. https://www.phnompenhpost.com/national/chinese-schools-back-brink.

Lee, Cher Leng, 'Grandmother's tongue: Decline of Teochew language in Singapore.' in W. Li (ed.), *Multilingualism in the Chinese diaspora worldwide: Transnational connections and local social realities* (London: Routledge, 2015): 196–215. https://doi.org/10.4324/9781315759371.

Lee, Hugo Yu-Hsiu, 'Losing Chinese as the first language in Thailand', *Asian Social Science*, 10/6 (2014): 176–93. https://doi.org/10.5539/ass.v10n6p176.

Li, Danni, *Acoustic analysis of the tones in the Shantou dialect* (MA thesis, University of Massachusetts Amherst, 2014). ScholarWorks@UMass Amherst. https://scholarworks.umass.edu/masters_theses_2/29.

Li, Wei, Saravanan, Vanithamani and Ng, Julia Lee Hoon, 'Language shift in the Teochew community in Singapore: A family domain analysis', *Journal of Multilingual and Multicultural Development*, 18/5 (1997): 364–84. https://doi.org/10.1080/01434639708666326.

Li, Yongming 李永明, 潮州方言 *Chaozhou dialect* (Hong Kong: 中華書局, 1959).

Li, Yongming 李永明, '新加坡潮州话的外语借词和特殊词语 Loanwords and special words in Singapore Teochew', 方言, 7 (1991): 56–63.

Lim, Alvin Cheng-Hin, *Desiring Cambodia* (unpublished PhD dissertation, University of Hawai'i at Mānoa, 2011).

Lim, Hiong Seng, *Handbook of the Swatow vernacular* (Singapore: Koh Yew Hean Press, 1886).

Lin, Lun-Lun 林伦伦, '潮汕方言声调研究 Research on the tones of Chaoshan dialect', 语文研究, 1 (1995), 52–9.

Lin, Lun-Lun 林伦伦, '澄海方言研究 *Research on the Chenghai dialect*.' (Shantou: 汕头大学出版社, 1996).

Lin, Lun-Lun 林伦伦, '潮汕方言和泰语的双向借词及其演变发展 Thai loanwords in Chaoshan dialect and Chaoshan dialect loanwords in Thai language', 民族語文, 2 (2006), 24–30.

Lin, Lun-Lun 林伦伦 and Chen, Zhao-Er 陈照儿, '外砂话声调与澄海话, 汕头话的比较研究 Tones of Waisha dialect and its comparative study with Chenghai dialect and Shantou dialect', 汕头大学学报 (人文社会科学版) *Shantou University Journal* (*Humanities & Social Sciences Bimonthly*), 27/3 (2011): 56–60.

Lin, Qing, *The diachrony of tone sandhi: Evidence from Southern Min Chinese* (Berlin: Springer, 2019). https://doi.org/10.1007/978-981-13-1939-6.

Low, Elissa Jia Min, *A sketch grammar of Singapore Teochew* (final year project, Nanyang Technological University, 2014). Digital Repository of NTU. https://hdl .handle.net/10356/65807.

Ly, Y., *Heaven becomes Hell: A survivor's story of life under the Khmer Rouge* (ed. J.S. Driscoll, Yale University Southeast Asia Studies, 2000).

Manomaivibool, Prapin, *A study of Sino-Thai lexical correspondences* (PhD dissertation, University of Washington, 1975). ResearchWorks Archive. http://hdl.handle .net/1773/11132.

Matthews, Stephen, Xu, Huiling and Yip, Virginia, 'Passive and unaccusative in the Jieyang dialect of Chaozhou', *Journal of East Asian Linguistics*, 14/4 (2005): 267–98.

Matthews, Stephen and Yip, Virginia, 'Passive, unaccusative and pretransitive constructions in Chaozhou,' in R. Djamouri, B. Meisterernst and R. Sybesma (eds.), *Chinese linguistics in Leipzig* (漢語語言學在萊比錫) (EHESS, Centre de Recherches Linguistiques sur l'Asie Orientale, 2008): 163–74.

McFarland, Joanna R., *A preliminary comparative analysis of Cambodian Teochew: Evidence for contact-induced change* (unpublished MA dissertation, University of Hong Kong, 2017).

McFarland, Joanna R., *A grammar of Cambodian Teochew* (PhD dissertation, Nanyang Technological University, forthcoming).

Miao, Ruiqin, *Loanword adaptation in Mandarin Chinese: Perceptual, phonological and sociolinguistic factors* (PhD dissertation, Stony Brook University, 2005). RUcore: Rutgers University Community Repository. https://rucore.libraries.rutgers.edu/ rutgers-lib/38301/.

Morita, Liang, 'Discussing assimilation and language shift among the Chinese in Thailand', *International journal of the sociology of language*, 186 (2007): 43–58. https://doi.org/10.1515/IJSL.2007.041.

Myers-Scotton, Carol, *Duelling languages: Grammatical structure in codeswitching* (Oxford: Oxford University Press, 1993).

Ng, Julia Lee Hoon, *Language maintenance and language shift in the Teochew Chinese community* (academic exercise, National Institute of Education, 1996). NIE Digital Repository. http://hdl.handle.net/10497/1723.

Nyíri, Pál, 'Investors, managers, brokers, and culture workers: How the "new" Chinese are changing the meaning of Chineseness in Cambodia', *Cross-Currents: East Asian History and Culture Review*, 1/2 (2012): 369–97. https://doi.org/10.1353/ach.2012.0012.

Nyíri, Pál, '"New migrants" from the PRC and the transformation of Chinese media: The case of Cambodia', in W. Sun and J. Sinclair (eds.), *Media and communication in the Chinese diaspora* (London: Routledge, 2015): 15–31. https://doi.org/10.4324/9781315717265.

Ong, Teresa Wai See, *Language maintenance in Malaysia: A case study of the Chinese community in Penang* (PhD thesis, Griffith University, 2018). Griffith Research Online. http://hdl.handle.net/10072/382738.

Pan, Jiayi 潘家懿, '柬埔寨第三代华人潮汕话记略 Cambodia's third generation Chaoshan speakers', in R. Li (ed.), 东南亚华人语言研究 *Studies of Chinese Languages in Southeast Asia* (Beijing: 北京语言文化大学出版社 Beijing Language and Culture University Press, 2000).

Panh, Rithy and Bataille, Christophe, *The elimination: A survivor of the Khmer Rouge confronts his past and the commandant of the killing fields* (trans. J. Cullen, New York City: Other Press, 2012).

Peng, Anne Elise, 'Head-final and head-initial relative clauses in Jambi Teochew', in K. Otaki, H. Takeyasu and S.-I. Tanigawa (eds.), *Online Proceedings of GLOW in Asia Workshop for Young Scholars* (2011): 262–76. http://faculty.human.mie-u.ac.jp/~glow_mie/Workshop_Proceedings/20Peng.pdf.

Peng, Anne Elise, *Aspects of the syntax of Indonesian Teochew* (unpublished PhD dissertation, University of Delaware, 2012).

Phadungsrisavas, Veerakit, *An acoustical comparative study of Swatow Chinese tones spoken in five regions of Thailand* (unpublished MA thesis, Mahidol University, 2008).

Pou, Saveros and Jenner, Philip N., 'Some Chinese loanwords in Khmer', *Journal of Oriental Studies*, 11/1 (1973): 1–90.

Quek, Huan Ting Geraldine, *Investigating language loss within Singapore's Teochew community* (unpublished final year project, Nanyang Technological University, 2013).

Rappa, Antonio, *Thailand's Chinese population: Teochiu speakers and political identity* (unpublished paper, 2014).

Sak-Humphry, Chhany, *Colloquial Cambodian: The complete course for beginners* (new edn., London: Routledge, 2015). https://doi.org/10.4324/9780203120651.

San Roque, Lila, Gawne, Lauren, Hoenigman, Darja, Miller, Julia Colleen, Rumsey, Alan, Spronck, Stef, Carroll, Alice and Evans, Nicholas, 'Getting the story straight: Language fieldwork using a narrative problem-solving task', *Language Documentation & Conservation*, 6 (2012): 135–74. http://hdl.handle.net/10125/4504.

Schliesinger, Joachim, *Ethnic groups of Cambodia, volume 3: Profile of Austro-Thai and Sinitic-speaking peoples*, vol. 3 (Banglamung: White Lotus, 2011).

Selinker, Larry, *Rediscovering interlanguage* (London: Routledge, 1992). https://doi.org/10.4324/9781315845685.

Shih Shu-mei, 'Introduction: What Is Sinophone Studies?', in Shih Shu-mei, Chien-Hsin Tsai and Brian Bernards (eds.), *Sinophone Studies: A Critical Reader* (New York: Columbia University Press, 2013): 1–16.

Sim, Tze Wei, 'Why are the native languages of the Chinese Malaysians in decline', *Journal of Taiwanese Vernacular*, 4/1 (2012): 62–95.

Siphat, Touch, 'Chinese capitalism in Cambodian sociopolitical contexts: The role of ethnic Chinese in the Cambodian economy,' in Y. Santasombat (ed.), *Chinese capitalism in Southeast Asia* (London: Palgrave Macmillan, 2017): 181–205. https://doi.org/10.1007/978-981-10-4696-4.

Stenberg, Josh, 'Multilingualism and the West Kalimantan Hakka,' in Li Wei (ed.), *Multilingualism in the Chinese diaspora worldwide: Transnational connections and local social realities* (London: Routledge, 2015): 123–40. https://doi.org/10.4324/9781315759371.

Stevens, Christine Audrey, *"New life in the freedom country": Young Cambodians in Adelaide* (PhD thesis, University of Adelaide, 1990). Adelaide Research & Scholarship. http://hdl.handle.net/2440/19370.

Swadesh, Morris, 'Towards greater accuracy in lexicostatistic dating', *International journal of American linguistics*, 21/2 (1955): 121–37. https://doi.org/10.1086/464321.

Tadmor, Uri, 'Loanwords in the world's languages: Findings and results', in M. Haspelmath and U. Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook* (Berlin: De Gruyter Mouton, 2009): 55–75. https://doi.org/10.1515/9783110218442.55.

Tan, Danielle, *La diaspora chinoise du Cambodge: Histoire d'une identité recomposée* (unpublished MA thesis, Institut d'Etudes Politiques de Paris, 2006).

TCKnow LLC., *WhatTCSay* (Version 1.30) [Mobile app], 2015. App Store. https://apps.apple.com/us/app/whattcsay/id550890802.

Tea, Van and Nov, Sokmady, 'The ethnic Chinese in Cambodia: Social integration and renaissance of identity,' in *Ethnic Groups in Cambodia* (Phnom Penh: Center for Advanced Study, 2009): 235–80.

Teo, Kok, *A sociolinguistic description of the Peranakan Chinese of Kelantan, Malaysia* (PhD dissertation, University of California, Berkeley, 1993). eScholarship. https://escholarship.org/uc/item/2hd9m92j.

The Teochew Store, *Top 100 surnames in Teochew* (accessed 27 March 2016). https://www.theteochewstore.org/blogs/latest/40690499-top-100-surnames-in-teochew-updated-27-march-2016.

Ung, Loung, *First they killed my father: A daughter of Cambodia remembers* (Phnom Penh: Random House, 2017).

Veniranda, Yohana, *Perfective aspect and negation in Pontianak Teochew* (PhD dissertation, University of Delaware, 2015). University of Delaware Library Institutional Repository. http://udspace.udel.edu/handle/19716/17687.

Veniranda, Yohana, 'Oral and nasal vowels in Pontianak Teochew', *LLT Journal: A Journal on Language and Language Teaching*, 18/2 (2016): 107–24.

Verver, Michiel, 'Templates of "Chineseness" and trajectories of Cambodian Chinese entrepreneurship in Phnom Penh', *Cross-Currents: East Asian History and Culture Review*, 1/2 (2012): 291–322. https://doi.org/10.1353/ach.2012.0017.

Verver, Michiel and Koning, Juliette, 'Toward a kinship perspective on entrepreneurship', *Entrepreneurship Theory and Practice*, 42/4 (2018): 631–66. https://doi.org/10.1177/1042258718783431.

Wang, Xiaomei, *Mandarin spread in Malaysia* (Kuala Lumpur: University of Malaya Press, 2012).

Wang, Xiaomei, 'The Chinese language in the Asian diaspora: A Malaysian experience', in G. Leitner, A. Hashim and H.-G. Wolf (eds.), *Communicating with Asia: The future of English as a global language* (Cambridge: Cambridge University Press, 2016): 205–15. https://doi.org/10.1017/CBO9781107477186.014.

Weinreich, Uriel, *Languages in contact: Findings and problems* (Berlin: De Gruyter Mouton, 1979). https://doi.org/10.1515/9783110802177.

Willmott, William E., *The Chinese in Cambodia* (Vancouver: University of British Columbia Press, 1967).

Willmott, William E., 'Cambodia', in L. Pan (ed.), *The Encyclopedia of the Chinese Overseas* (2nd ed., Singapore: Editions Didier Millet, 2006): 144–50.

Xu, Huiling, *Aspect of Chaozhou grammar: A synchronic description of the Jieyang variety* (Journal of Chinese Linguistics Monograph Series 22, 2007).

Xu, Huiling and Matthews, Stephen, 'On the polyfunctionality and grammaticalization of the morpheme *kai* in the Chaozhou dialect,' in F.H. Yap, K. Grunow-Hårsta and J. Wrona (eds.), *Nominalization in Asian Languages: Diachronic and typological perspectives* (Amsterdam: John Benjamins, 2011): 109–24. https://doi.org/10.1075/tsl.96.03xu.

Xu, Huiling and Matthews, Stephen, 'On the semantic continuum of the causative constructions in Chaozhou and Taiwanese Southern Min', in *Overseas diaspora and homeland culture: proceedings of selected works of the ninth International Chaozhou Studies Symposium* (Guangdong: Huacheng Publishing House, 2013): 232–44.

Xu, Yuhang 徐宇航, '潮州方言百余年来韵母演变的研究 Phonological changes of vowels in Chaozhou dialect: From mid-nineteenth to early twenty-first century', 语言学论丛, 53 (2016): 241–60.

Yeo, Pamela Yu Hui, *A sketch grammar of Singapore Teochew* (final year project, Nanyang Technological University, 2011). Digital Repository of NTU. https://hdl .handle.net/10356/93949.

Zhou, Daguan, *A record of Cambodia: The land and its people* (trans. P. Harris, Chiang Mai: Silkworm Books, 2007).

The Nature of Sinitic Lexicon in Bazaar Malay and Baba Malay in Singapore

Khin Khin Aye

1   Introduction

Voluntary or involuntary, human migration is fundamental to Asian history and to the world in general. Wherever people migrate, they introduce their languages, cultures, religions and beliefs to the host communities. Chinese emigration to other countries started as early as 1000 CE. From the mid-nineteenth century, their migration to other countries has resulted in more than 30 million so-called "Overseas Chinese", who live outside mainland China and include over 20 million people in Southeast Asia (Lockard 2013), affecting the world's demographics, economy, culture, and language, just to name a few. In terms of population, according to Hay (2008), ethnic Chinese range between 0.8% and 76% of the population of each Southeast Asian nation. These numbers do not take into consideration partially assimilated Chinese in these communities. Ethnic Chinese make up 34% of the population (6 million) in Malaysia and 76% (2 million) in Singapore, the majority of whom are descendants of Hokkien speakers present since the early days of Singapore. Economic dominance of the Chinese is felt in countries such as Malaysia, Singapore, Thailand, and Indonesia. As pointed out by Hays (2008), the economies are mostly controlled by rich Chinese in these countries.

Migration also set the stage for immigrant Chinese and host communities, who spoke a range of different languages, to communicate and interact. This required the people involved to tap into the full repertoire of languages at their disposal. As noted by the mainland Chinese novelist Wang Anyi 王安憶 in the context of Singapore and Malaysia, but equally valid for other Southeast Asian countries, "Hua people must endure speaking the language of another ethnic group" (Ng 2013: 84). Explicit or implicit lexical borrowing between the languages involved in such processes is one of the most common consequences of contact, and has been observed in all world languages (Hoffer 2005). The Loanword Typology Project, coordinated by Haspelmath and Tadmor between 2004 and 2008, looked at loanwords in 41 world languages. Based on these findings, Tadmor (2009) ranks English among the "high borrowers" with 40%

of its total words identified as loanwords. Mandarin hit the study's bottom of "low borrowers" with only 1.2%. Evidence furthermore shows that loanwords from Chinese varieties have been found, for example, in Indonesian and Malay. Jones (2009: 9) identifies the word *tahu* 豆腐 'bean curd' as the region's earliest Hokkien loanword, found in a tenth-century Old Javanese inscription. The absence of similar evidence in Old Malay inscriptions led him to assert that Chinese loanwords entered the Malay language only during the Qing Dynasty (1664–1912). Scholars like Jones (2009), Pou & Jenner (1973), and Egerod (1959) have investigated Chinese loanwords in, respectively, Malay, Khmer, and Thai. Yet compared to the large body of scholarship on loanwords from English in other world languages or foreign loanwords into English, for example Durkin (2014), studies on Sinitic loanwords into other languages have been relatively limited.

Another migration-related outcome is the emergence of lingua francas: contact languages used for communication among people of diverse backgrounds. Two cases in point are Bazaar Malay and Baba Malay, both of which are Malay-based contact varieties with different sociolinguistic statuses. The former is a Malay-lexified lingua franca used for several centuries in the multilingual trade ports of Southeast Asia. It existed in Singapore since the early days of its modern history, beginning in 1819. As pointed out by Collins (1984), it was nobody's first language, yet it was the only form of Malay spoken by elderly non-Malay Singaporeans and Malaysians. Singapore Bazaar Malay was used by Malays to non-Malays, non-Malays to Malays and Baba Chinese, among non-Malays when they did not have any common language, and among Chinese, Malays, or Indians who spoke mutually comprehensible languages (Aye 2006: 24). Baba Malay, on the other hand, is the first language of the Peranakan Chinese community in Singapore and Malaysia. Peranakan Chinese are Straits-born Chinese and their descendants, who have developed a rich creolized culture, language, and literature. This was the community of Chinese in Singapore and Malaysia that pioneered romanized Baba Malay publications, which were translated mostly from Chinese and to a lesser extent from English literature. Researchers such as Tan (1993: 40) and Ansaldo and Matthews (1999: 45) zero in on the connection between these two Malay-based varieties, noting their shared Chinese input and the possible contribution of Batavia Bazaar Malay to the formation of Baba Malay (cf. Pakir 1986). Pakir (1986) and Aye (2006), among others, shed light on Sinitic influence on Singapore Baba Malay and Bazaar Malay, though their studies focus more on describing linguistic features as a whole while also considering their socio-cultural and sociolinguistic contexts. The historical and demographic accounts of modern Singapore

analysed in these studies point to a shared linguistic ecology in which both varieties emerged, featuring Hokkien as their major substrate.

As a Myanmar Chinese with Fuzhou as my first language, married into a Hokkien family in Myanmar, I have long been motivated to investigate Sinitic influences on the languages and cultures of their host communities. My linguistic and academic background started to converge during my postgraduate diploma course, when I started to study and write about Hokkien kinship terms and code-switching in a Sino-Myanmar family. Against the backdrop of limited scholarship on Sinitic loanwords (except the publications mentioned previously) measured against the ubiquity of Chinese people in Southeast Asia, this chapter investigates the nature of the Hokkien lexicon in Bazaar Malay and Baba Malay used in Singapore. The Bazaar Malay data have been collected from ten consultants between 2002 and 2004, whereas the Baba Malay data were acquired through a textual analysis of the novel *Si Hitam Yang Chantek*, the Baba Malay version of Anna Sewell's (1797–1884) famous *Black Beauty*. It was one of very few translations into Baba Malay not based on a Chinese original. Published by Singapore Methodist Publishing House in 1913 and translated by Goh Hood Keng 吳佛經, the book was intended for Christian Peranakan Chinese (Proudfoot 1993: 259). It was selected as the use of uncommon Hokkien cultural terms in this book was minimal compared to Baba Malay publications translated from Chinese stories (Shellabear 1919: 383). As such, the idiom in *Si Hitam Yang Chantek* is closer to the spoken language of that period.

It is well-known in the field of Pidgin and Creole Studies that substrate languages and their speakers are generally stigmatized (Bao 2003). This chapter argues that extralinguistic factors – the numerical dominance of Hokkien speakers in the history of Singapore's demographics and their spread in all walks of life – overruled this stereotype and facilitated the adoption of Hokkien lexicon into these Malay varieties. Differences in their sociolinguistic functions in Malaysia and Singapore have determined the degree of borrowing into both varieties.

2   Background

According to Wang (1991), Chinese emigration over the last two centuries features four major models with distinctive professions, geographical origins, distributions, and temporal dimensions: (1) the *Huashang* 華商 'Chinese traders', especially in Southeast Asia before 1850, mainly consisting of male traders who later intermarried with the local communities once settled down; (2) the

Huagong 華工 'Chinese coolies', consisting of male peasants, especially to North America and Australia between 1840s and 1920s as workers involved in mining, railway building, and agriculture sectors; (3) the *Huaqiao* 華僑 'Overseas Chinese', especially since 1911, by well educated professionals who were vital to the Chinese education of the descendants of earlier immigrants in Southeast Asia; and (4) the *Huayi* 華裔 'people of Chinese descent', that is, descendants of Chinese migrants from one foreign country (e.g. in Southeast Asia) to another (e.g. in Western Europe), especially since the 1950s, when their participation in economic sectors was curtailed (Pan 1990: 226).

Chinese emigration to Singapore and West Malaysia (formerly known as Malaya) reflects three main patterns of migration: (1) the *Huashang* migration, which led to the birth of the Peranakan or Straits-born Chinese community whose first language was Baba Malay; (2) the *Huagong* migration that led to later waves of Chinese migration into Singapore since 1824; and (3) the *Huaqiao* migration which between the 1920s and 1950s played a pivotal role in teaching Chinese to Chinese-descended Southeast Asians who had already settled down in the local communities (Pan 1990: 206; Poston and Luo 2007). Chinese in Southeast Asia came predominantly from the southern Chinese coastal provinces of Fujian and Guangdong and some from Hainan in southern China (Freedman 1960: 26; Hay 2008; Pyau Ling 1912: 75). Chinese people, speaking a number of Sinitic varieties, mostly live in Southeast Asia's urban areas, including a number of "Chinatowns", while others live in the rural areas (cf. Hay 2008; Poston and Wong 2016: 367).

Singapore as a multi-ethnic and multicultural island state is unique in a broader Southeast Asian context for having English (the language of business and administration), Mandarin, Malay (the National Language), and Tamil as its four official languages. Most Singaporeans are bilingual (or multilingual) and speak their "mother tongue" and English with varying degrees of fluency. According to Population Trends, out of a total population of 3,965,796 in 2017, the ethnic Chinese made up 74.3%, Malays (including those of Indonesian descent) 13.4%, Indians of different ethnic or linguistic groups 9%, and others – which include Eurasians and Europeans – 3.2%. In this linguistic ecology, Singapore Bazaar Malay (the lingua franca of this multilingual community) and Baba Malay (the first language of the Straits-born Chinese) historically existed together with the languages spoken by these four major groups. Given the massive language shift in Singapore to English, both Malay varieties now suffer the same fate of language endangerment. The position of the nation's official Malay variety, conversely, remains stable, although it has become somewhat uncommon outside the nation's Malay community and government announcements.

Bazaar Malay, also known as "Pasar Malay", "Pasar Melayu" or "Melayu Pasar", served for several centuries as a lingua franca in Southeast Asia's major commercial centres, including for nearly two centuries in Singapore (Collins 1987; van Minde 1997: 8; Paauw 2008).[1] The earliest trading contacts with India and China (Baxter 1985: 14; Prentice 1992: 374) and the associated commercial networks ultimately gave rise to several trade centres, such as Malacca on the southwest coast of the Malay Peninsula (before 1511) and later Singapore (since 1819). Business transactions in these trade centres were highly diverse, involving people from different origins, such as Chinese, Gujaratis, South Indians, Javanese, and Malays (Hall 1985: 86). This resulted in the development of Bazaar Malay to meet the communicative needs of those involved in the business transactions of these marketplaces. Its lexifier language was Malay, which was conventional given the latter's geographical distribution along the Straits of Malacca and its significant role in the region's administration, religion, culture, and commerce. Bazaar Malay is no one's first language, and at the time when I collected data in Singapore, its speakers (fewer than 1,000) mostly belonged to the older generations and some middle-aged workers whose daily routine required them to deal with elderly Singaporeans who did not speak English. Its role as the lingua franca had been replaced by English due to education policies and language campaigns launched in Singapore and resulting in a massive language shift (Poedjosoedarmo 1997).

Baba Malay, conversely, developed out of intermarriages between early Chinese immigrants and local speakers of Malay. It became the first language of the Peranakan or Straits-born Chinese and their descendants.[2] The forebears of the Baba Chinese were early Hokkien-speaking immigrants from Fujian who migrated to Malacca around the fifteenth century, long before the establishment of the British Straits Settlements (Purcell 1948; Vaughan 1971; Ansaldo & Matthew 1999: 39). Their Malay-lexified creole was a hybrid between Malay and Hokkien, which was comprehensible to native Malay speakers (Pakir 1986; Thurgood 1998; Lee 2014). These Straits-born Chinese also became pioneers in learning – and shifting to – English. Their knowledge of English, Hokkien, and Malay enabled them to act as go-betweens for business transactions between European firms and local trade centres (Platt and Weber 1980: 3). Unlike Bazaar Malay, Baba Malay had a well-established speech

---

1   Note that not all researchers may refer to precisely the same variety. As discussed in Paauw (2008: 9), the term "Bazaar Malay" is used to refer to two different things: (1) the contact language that emerged in the multi-ethnic West Malaysia (Adelaar 1991; Bakker 2003, 2004), and (2) the vehicular trade language that spread throughout the Indonesian islands (Adelaar and Prentice 1996).

2   The men of this community are known as *baba*s and women as *nyonya*s.

community with a rich culture, language, and literature. The hybrid nature of this community (Chinese, Malay, and English) is reflected in their dress, unique architectural style, culinary skills, and mixed language and literature. The uniqueness of the Baba Malay language, as noted by previous researchers, includes the use of (Hokkien) Chinese words and expressions for specific items of clothing, culture, cooking, daily life, etc. Its phonology and morphosyntactic properties that differ from other Malay dialects have received academic attention (Pakir 1986; Tan 1998, 2004; Lee 2014). Researchers such as Shellabear (1913), Png (1963), and Gwee (2006) have conducted Baba Malay lexical studies, while Gwee (1993) added phrases, sayings, and expressions. The Straits-born Chinese were the first Chinese in the Malay Peninsula to use a romanized form of their spoken language in their literary publications. While there is an ongoing debate on whether to classify Baba Malay as a creole (Lee 2014, for example) or a Malay dialect (Pakir 1986; Tan 1998), this chapter approaches it as a Malay-based contact language. Like Bazaar Malay, Baba Malay is fast disappearing in Singapore and Malaysia, with the increasingly extensive use of English among the younger generation. According to Lee (2014: vii), it has less than 1000 speakers in Singapore and Malacca (Malaysia) respectively.

Since the establishment of modern Singapore in 1819, its population has been strikingly multiracial, multicultural, and multilingual. The island's free trade, job openings, law and orderliness attracted immigrants from diverse ethnolinguistic and cultural backgrounds. This in turn offered an excellent breeding ground for contact varieties such as Bazaar Malay to evolve. Languages spoken by immigrants can be considered as input languages in the development of Bazaar Malay. They include various languages spoken by Malay, Chinese, Indian, Eurasian, and European people. In the development of Baba Malay, Malay and Chinese varieties, specifically Hokkien, played the most important role. Other Chinese immigrants were speakers of Teochew or other Min varieties, Cantonese, Hainanese, and Hakka. As the name suggests, Bazaar Malay has Malay as its lexifier and other languages as substrate elements. Baba Malay, on the other hand, has Malay as lexifier and Hokkien as its sole substrate. Unlike other parts of the region, where Chinese are a minority group, in Singapore the ethnic Chinese have since the 1840s predominated in terms of population and in most sectors of the economy, as trade was largely in the hands of Chinese businessmen and Europeans (Platt and Weber 1980: 3).[3]

---

3 Chinese makes up 50% of the total population in 1840, 61.2% in 1860, 67.1% in 1891, 72.4% in 1911, 75.1 in 1931, 76.9% in 1957 and 1980, 76.8% in 2000 (Aye 2013: 87; Bao 2001: 281; Population 2000 Census).

Furthermore, Chinese immigrants (excluding the Straits-born Chinese) engaged in all sorts of non-commercial activities (Tarling 1992: 10). As mentioned previously, Hokkien speakers constitute the majority of ethnic Chinese in Singapore and Hokkien has been used for Chinese inter-group communication; while the island's other Sinitic varieties exhibit a shared core grammar, most are mutually unintelligible (Kuo 1976: 11; Bao 2005). The economic and social dominance of Hokkien speakers helps us to explain the status of Hokkien as the major substrate language of Bazaar Malay and Baba Malay. This chapter's focus on the nature of Hokkien borrowing is likely to also contribute to a better understanding of the unique features of Singapore English and Malaysian English, although they fall outside its scope. All these varieties (Bazaar Malay, Baba Malay, Singapore English, and Malaysian English) have developed in roughly the same linguistic ecology and their typological similarities attest to this shared history.

3   Theoretical Framework and Methodology

As I will argue in this section, the theories of borrowing and relexification described in the work of Claire Lefebvre can best explain the existence of Hokkien lexicon in Baba Malay and Bazaar Malay. Both processes are at play whenever speakers of different languages are in contact. Information on borrowing and relexification is provided below, followed by a brief note on my research methodology.

3.1   *Borrowing*

Borrowing is the linguistic strategy adopted by speakers to discuss new entities or concepts in another language. It is a discernible outcome of contact with people coming from diverse linguistic, cultural, and religious backgrounds. According to Hockett (1997, cf. Hoffer 2005), four borrowing options are available to the speakers of a language:

1.   Loanwords: the adoption of words from a donor language, integrated into the borrowing language's grammatical system
2.   Loan shift: attributing native words with new meanings
3.   Loan translation or *calque*: the lexification of words in the borrowing language with syntactic and sematic properties from the donor language
4.   Loan blend: the use of a combination of a loanword and a native word

My analysis of Bazaar Malay and Baba Malay data suggests that all four options of borrowing can be observed in these varieties, though the quantity of each option differs. More details are provided and discussed in sections 4 and 5.

3.2   *Relexification*

In order to satisfactorily explain the nature of Hokkien lexicon in the two
Malay varieties, I use the concept of relexification as a mental process, as pro-
posed by Lefevre (1998), in tandem with two constraints put forward by Bao
(2005). According to Bao (2015: 15) and Aye (2006: 194), new lexical entries in
contact languages are adopted as a result of collaborative contributions from a
superstrate or lexifier language and one or more substrate languages.

Relexification was first defined by Muysken (1981: 61). According to Lefevre
(1998: 16), this process constructs new lexical entries in emerging contact
languages in two phases: (i) copying the substrate lexicon, and (ii) replacing
the phonological forms of substrate words with those from the lexifier lan-
guage, which is termed "relabelling" (Lefebvre & Lumsden 1994). According
to Lefebvre (1998: 17), while all lexical categories can be copied, relabelling is
semantically determined. In other words, this process takes place on the con-
dition that there is an overlap between the semantics of the lexical item in
the substrate and in the lexifier. The outcome of this process is a new lexical
entry in the lexifier language's phonological form, though with the substrate
language's semantic and syntactic properties.

Given the failure of this relexification process to yield some expected
Chinese aspectual categories in Singapore English, Bao (2005) proposes two
constraints to relexification – "system transfer" and "lexifier filter" – and identi-
fies specific and interdependent roles played by the substrate and the lexifier
language. A system transfer requires the compulsory transfer of the substrate's
entire grammatical subsystem, while a lexical filter only allows for the transfer
of the system which complies with the lexifier's surface structure require-
ments. In the discussion in section 5.2 on loan translations, we will see how
this is relevant for the Hokkien-influenced Malay varieties studied here.

3.3   *Methodology*

To identify Hokkien loanwords in the two contact varieties, I started by ana-
lysing Bazaar Malay data collected from ten consultants: four Chinese, four
Indians, one Boyanese, and one Malay. I met most of them between 2002
and 2005 in the Kolam Ayer Community Centre in the Geylang Bahru area of
Singapore. Closer socialization among the older residents of different races in
this area – along with the somewhat higher concentration of Malays, Indians,
and other minority groups[4] – may have caused these people to use Bazaar
Malay as a communicative bridge within this community, despite its declining

---

4   The population ratio of the Kolam Ayer Constituency at that time was: Chinese 77.81%,
     Malays 10.18%, Indians 10.84%, and others 1.17%.

use at that time. Most of them were born in Singapore and the rest had been living in Singapore for more than thirty years. The most salient characteristic of these consultants was that they had picked up Malay through informal acquisition. This was even the case with those who had learnt Malay formally, as they only learnt very basic vocabulary, although the duration of their education ranged from one to three years. Bearing in mind that triangulation is vital for valid fieldwork findings, data on my consultants' "relaxed", "natural" usage were collected by means of (i) survey questionnaires which I asked them personally, (ii) direct elicitation via a Translation Test, (iii) indirect elicitation through different means, and (iv) other strategies whenever appropriate (see Gil 2001: 121; Mithun 2001: 35; Milroy and Gordon 2003: 58).[5]

For Baba Malay, I looked at written rather than elicited data. According to Yoong & Zainab (2002: 1), in order to accommodate the reading needs of the Baba community and their desire to have a written language, newspapers, poems, novels, magazines, and translated Chinese and English stories were published between the late 1880s and 1950 in romanized Baba Malay. Some examples of newspapers include the *Straits Chinese Herald* (*Surat Khabar Peranakan*) and *Bintang Timor* from 1894 and *Khabar Slalu* (*Daily News*) from 1924. In 1930, the first Baba Malay weekly, *Bintang Pranakan*, was published in romanized Malay. Yoong and Zainab (2002) found 68 titles of translated Baba Malay literature from Chinese stories between 1889 and 1950.[6] One of the most famous examples of these works is *Sam Kok* 三國 (1892–96) translated from Luo Guanzhong's 羅貫中 *Romance of the Three Kingdoms* by Chan Kim Boon 曾錦文 (1851–1920). Ian Proudfoot's list of early Malay printed books until 1920 contains 16 entries of books intended for Christian readers, published by the American Mission Press and Methodist Publishing Press in Singapore. Among them we find Baba Malay translations of English novels, such as *Kmnangan Miriam Kristofer* (The Victory of Miriam Christopher by H.R. Calkins) translated by Mrs. J.M. Hoover (Proudfoot 1993: 306).

---

5   As my consultants were elderly Singaporeans, I raised questions close to their heart to grab their attention and trigger their reaction. These included questions about a popular lottery (*toto*) and childhood experiences in the case of a male consultant and about former colleagues in school and raising children in the case of a female consultant. Doing so guaranteed their emotional involvement and I managed to get a lot of spontaneous speech data. I also joined some of the activities of my consultants, including playing gate ball with a recorder hanging around my neck, and chatted informally with this entire group of senior citizens.

6   18 works between 1889 and 1909; 10 between 1910 and 1919; 39 between 1935 and 1939; and 1 in 1950.

For this chapter, I have chosen *Si Hitam Yang Chantek*, which was translated by the aforementioned Goh Hood Keng and published by Methodist Publishing House in 1913 (Proudfoot 1993: 681). The book has 298 pages and consists of five parts divided into 45 chapters. The English original, *The Black Beauty*, was written by Anna Sewell. This first-person narrative is the autobiography of a horse named "Black Beauty". Using Antconc concordance software, I have conducted a textual analysis of the novel in 2005 and 2006, inspired by Pakir (1986) and subsequently filtered through the data of Lee (2014).[7]

4　　　The Nature of Hokkien Lexicon in Bazaar Malay and Baba Malay

While the majority of the lexicon of Bazaar Malay and Baba Malay is Malay-derived, the influence of its Hokkien substrate is evident in both languages, although its pervasiveness differs across these varieties, in line with their different sociolinguistic statuses in Singapore.

Out of 1073 words in the glossary of Baba Malay provided by Lee (2014), Hokkien loanwords make up 15.56% (167 words / expressions), followed in quantity by some English loanwords. A total of 219 Hokkien words were identified by Pakir (1986: 116) in her Baba Malay data. These are categorized into 12 categories, only four of which were found in Bazaar Malay. These four include:

1) Terms of address, including kinship terms and personal pronouns,
2) Temporal expressions,
3) Business terms, and
4) Culinary terms.

In what follows, these different categories of Hokkien lexicon in Bazaar Malay and Baba Malay are described in detail, paying particular attention to kinship terms and personal pronouns.

4.1　　*Kinship Terms*

The Hokkien lexical substrate in the use of terms of address in Bazaar Malay and Baba Malay has frequently been mentioned in the literature on contact linguistics and Malay linguistics (Baxter 1985; Pakir 1984; Holm 1989: 578–79; Adelaar & Prentice 1996: 675; among many others). In Baba Malay, Pakir (1986: 107) lists Hokkien kinship terms in use among elderly Baba Chinese, who tend to have

---

7　Professor Bao Ziming kindly allowed me to use these data which I collected and analysed when working as a student research assistant at the Department of English Language and Literature, National University of Singapore between 2005 and 2006.

a stronger sense of kinship and familial hierarchy. Out of the 72 kinship terms in Baba Malay provided by Lee (2014: 401–2), 55 (71.4%) are Hokkien-derived, including *gong* 公 or *gong gong* 'grandfather', *chau* 'elder sister's husband',[8] etc. and the remaining 17 are from Malay. In the novel *Si Hitam Yang Cantik*, the use of the Hokkien word *ngkong* 'grandpa' 俺公 is found 18 times and *tachi* 大姐 'elder sister' three times. Consider the following example from Chapter 4:

1) | *Saya* | *banyak* | *terima* | *kasih* | *kerana* | *lu* | *punya* | *baik* |
   |--------|----------|----------|---------|----------|------|---------|--------|
   | 1SG | many | accept | love | because | 2SG | punya | good |

kepada	*tachi*	*saya.*
to	elder.sister	1SG

   'I thank you very much for your good intention to my elder sister.'

It is not surprising to find a much smaller usage of such terms in Bazaar Malay, as this is a variety not just belonging to ethnic Chinese (in contrast to Baba Malay). Yet even Bazaar Malay speakers have been observed to use the appropriate kinship terms from Malay, Tamil, Hokkien – such as *angso* 俺嫂 'sister-in-law' to address a Chinese female colleague – and English (*uncle* and *aunty*) to show respect, intimacy, and friendliness among interlocutors (see also Collins 1987: 164).

4.2   *Personal Pronouns*

In both Bazaar Malay and Baba May, the use of Hokkien first personal pronoun *gua ~ wa* 我 and the second personal pronoun *lu* 汝 can be observed (see also Lee 2014: 396). While the direct use of these personal pronouns is restricted to Baba Malay according to Pakir (1986: 106), my study demonstrates otherwise.[9] It is not uncommon in the marketplace to hear an Indian man address a Chinese shopkeeper with *lu* 'you' and refer to himself as *wa* 'I'. Although the use of Malay *awak* 'you' and *saya* 'I' often marked the beginning of such conversations, *lu* 'you' and *wa* 'I' quickly became the norm as they went on. This widespread phenomenon substantiates the status of *gua ~ wa* and *lu* as part of the Bazaar Malay lexicon. Table 4.1 illustrates Hokkien influence on Bazaar Malay and Baba Malay personal pronouns.

---

8 Presumably from Hokkien *chiá-hu* 姐夫 (Gwee 2006: 54).

9 Both are also common in Jakarta Indonesian (Adelaar 1991: 26), whereas *lu* as a 2SG is also observed in Kupang Malay (Paauw 2008: 166, 463).

TABLE 4.1   Personal pronouns in Malay, Hokkien, Baba Malay, and Bazaar Malay

Language	First person		Second person		Third person	
	Singular	Plural	Singular	Plural	Singular	Plural
Malay	saya, aku	kita (INC), kami (EXC)	awak, kamu, kau, anda	awak, semua	dia ~ ia	mereka
Hokkien	gua ~ wa 我	lan 咱 (INC), gun ~ guan 阮 (EXC)	lu, li 汝	lin 恁	i 伊	in 個, i-lang ~ i~nang 伊儂
Baba Malay	saya, gua	kita	lu	luorang	dia	diaorang
Bazaar Malay	saya, aku, wa	kitaorang, kita	awak, lu	luorang	dia	diaorang ~ diorang

ADAPTED FROM PAKIR (1986), LEE (2014: 30), AND AYE (2006: 72)

Note that the use of *gua* is considered impolite by Malays and even by some Chinese in my Bazaar Malay study. In the novel *Si Hitam Yang Chantek*, the perception of *gua* as rude is reflected in its frequency of occurrence – *gua* 49 times vs. and *saya* 2522 times – although it may also be due to the social status of *Black Beauty*, with its narrator horse as the protagonist. Yet, as commented by my participants, the direct adoption of Hokkien pronouns in Bazaar Malay manifests a close affinity between interlocutors of different ethnicities.

4.3   *Temporal Expressions*

In Baba Malay, temporal expressions of Hokkien origins are related to culture and in particular the days of special occasions marked by the lunar calendar. Lee (2014) lists ten Hokkien words indicating such days, including *chay-it* 初一 'the new year's day' and *chap-gor* 十五 'the fifteenth day of lunar month'. I did not detect any such terms in my Bazaar Malay data, as these Hokkien terms portray a specific Peranakan Chinese identity.

4.4   *Business Terms*

Some Bazaar Malay terms dealing with business are Hokkien words. My data show two such terms: *kongsi* 公司 'a company or firm; to share' and *tau ke* 頭家

'a well-to-do Chinese; an employer; the head of the business'.[10] Nine instances of *tau ke* were also observed in the Baba Malay novel (Chapter 38). See below for an example:

2) *Tapi   betul   ada   juga   tauke-tauke   yang   murah   hati*
   but   true   have   also   RED-employer   REL   kind   heart
   kepada   kita orang …
   to   1PL
   'But, it is true that there are also employers who are kind to us'.

With reference to the term *tauke*, Pakir (1986: 114) considers Malay as the source of this term. However, I would argue that it comes from Hokkien, even though it is now part of the Malay lexicon. In Hokkien, the term *tauke*, which literally means 'first family', is a compound made up of *tau* 頭 'head' and *ke* 家 'family'. It is also used among Hokkien speakers as well as non-Chinese Burmese in Myanmar, where Malay is not spoken.

4.5   *Culinary Terms*

Hokkien terms for cooking ingredients and dishes identified in Bazaar Malay include *cincau* 清草 'grass-jelly', *mi* 麵 'yellow noodles', *mihun* 麵粉 'rice vermicelli', *twahun* 大粉 'bigger rice vermicelli', *mi siam* 米暹 'sweet-and-sour noodles', *popiah* 薄餅 'spring rolls', *tau cio* 豆醬 'preserved soy beans', *tauge* 豆芽 'bean sprouts', *taukwan* 豆乾 'bean curd', and *tau yu* 豆油 'soy sauce'. This list mostly comes from a consultant who was preparing food in her kitchen while explaining to me how to do so. Hokkien words and phrases in this category are not exhaustive and there are many more of them in Bazaar Malay. These words are also found in Baba Malay. The following is an example from *Si Hitam Yang Chantek* (Chapter 44):

3) Kadang-kadang   dia   kasi   saya   makan   bangkuang,   atau
   sometimes   3SG   give   1SG   eat   turnip   or
   berdiri   di tepi   saya
   stand   near   1SG
   'Sometimes, he fed me turnip or stood near me.'

Note that the word *bangkuang* is also observed in Lee's data (2014), where its origin is indicated as Hokkien. Two points merit our attention: the referent of *bangkuang* (or *mangkuang*), and the question of its etymological origins.

---

10   The meanings of *kongsi* and *tau ke* are as given in Pakir (1986: 114).

Though it is glossed as 'turnip', following Lee (2014), a more appropriate gloss would be 'jicama (*Pachyrhizus erosus*)' – also known as 'sweet turnip' or 'Chinese turnip' – which is one of the ingredients of Hokkien *popiah* 薄餅 (see also malaysiavegetarianfood.com). Some people argue it is of Malay origin, which I do not believe to be accurate. In Standard Malay, the product is referred to as *ubi sengkuang*, not *bangkuang*. Lee (2014: 471) also lists it as a Hokkien word. In Burmese Hokkien, turnip is *tuar tau chai* 大頭菜, jicama is *kua jiu* 芥薯, and sweet potato is *han jiu* 番薯. The jicama is known as *tau aw jiu* 豆仔薯 in Taiwanese Hokkien, whereas in Singapore, *sa koh* 沙葛 and *bangkuang* 芒光 are common names. Hence, the word *bangkuang* strikes me as a Hokkien loanword unique to the Malay-speaking regions.

As these Hokkien-derived culinary terms can be heard in multiple contexts in Singapore, it is logical to conclude that they are part of the common core of the Singaporean linguistic repertoire, regardless of ethnic backgrounds. The same process is observed in Singapore Colloquial English. As these terms are also observed in other colloquial varieties of Malay, Pakir (1986) views them as Malay components which may have entered Baba Malay via Malay. As loanwords from different languages have been adopted into Malay via religion – such as Arabic words through Islam and Sanskrit words through Hinduism and Buddhism – the existence of Hokkien loanwords in generic Malay, as argued by Pakir (1986), is not unique to Baba Malay. Yet I agree with Ansaldo and Mathews (1999: 50) that these words may well have "*entered* [their emphasis] Malay through Hokkien speakers shifting to Malay and subsequently spread" to other Malay dialects, as there have been contacts between China and Southeast Asia through envoy, Buddhist pilgrims, and traders and this is hardly a recent phenomenon (Tan 1988: 28). Rather than being the recipient, we may therefore envision Baba Malay as the donor of these words.

5   Loan Shifts, Loan Translations, and Loan Blends

In what follows, I will pay attention to three common manifestations of language contact and lexical borrowing – loan shifts, loan translations, and loan blends – providing examples in each category of Hokkien influence on the two Malay varieties being researched here.

5.1   *Loan Shifts*

As discussed previously, some Malay-derived words in Baba Malay display non-mainstream meanings. One prominent example is found in the religious

sphere. In contemporary Malay, a number of religious terms have been claimed as the prerogative of specific ethno-religious groups. For example, the term *Allah* – derived from Arabic – was ruled in 2013 as a Muslim-only word by a Malaysian Court, inviting criticism from "Muslim thinkers and groups around the world" (Mandal 2016: 390; see also Leow 2016: 218–19). In multicultural, multilinguistic, and multi-ethnic countries like Singapore or Malaysia, the legal restriction of words to certain groups is untenable on a linguistic level, among many things, due to the aforementioned proclivity towards loan shifts, which commonly emerge from contact situations. We recall here that a loan shift refers to the adaptation of a word from the lexifier language, in this case Malay, to a shifted meaning. Thus, in the novel *Si Hitam Yang Cantek*, intended for the Christian Peranakan community, 'God' is referred to as *Tuhan Allah* (21 times) and *Allah* (4 times), compared its more neutral equivalent *Tuhan* (1 time). The higher frequency of *Tuhan Allah* or *Allah* substantiates that Christians commonly used this word. However, no such usage was found in this study's Bazaar Malay data, presumably because of its use in limited contexts.

5.2   *Loan Translations*

Loan translations or calques, which are analysed by Lefebvre (1998) through the prism of relexification, can be observed in the pluralization of the second and third personal pronouns and the formation of possessive pronouns. This process encompasses copying Hokkien semantic and morphosyntactic properties and relabelling them in Malay. Calques like these are commonly observed in contact situations (Sebba 1997: 92, 119).

Both in Baba Malay and Bazaar Malay, the plural forms of the personal pronouns are formed by adding *orang* 'person; people' to the respective pronouns:

SG	+ *orang* '**person**'
lu	*luorang* 'you (PL)'
dia	*diaorang* 'they'

Generic Malay exhibits personal pronouns with special plural forms, e.g. *kita* (1PL.INC) or *kami* (1PL.EXL) 'we' and *mereka* (3PL) 'they'. Yet in *Si Hitam yang Chantek*, 344 instances of *diaorang* 'they' and no single instance of the Malay *mereka* were found.

Plural forms in Hokkien, by contrast, are derived by adding *lang ~ nang* 儂 'person; people' or its contracted form *-n* immediately after the respective pronouns (Bodman 1955: 82; Ansaldo & Matthews 1999: 31):

SG	+ *lang* ~ *nang* / *-n* 'person'
gua 我 'I'	*guan* ~ *gun* 阮 'we (EXC)'
li 汝 'you (SG)'	*lin* 恁 'you (PL)'
i 伊 's/he'	*in* 個 / *i-lang* ~ *i-nang* 伊儂 'they'

It follows from this that the Hokkien use of *lang* / *-n* 'person' as plural marker was relabelled in Bazaar Malay and Baba Malay – using the inherited noun *orang* 'person; people' – and added to singular pronouns to derive their plural forms. The outcomes of this relexification process or calque of Hokkien plural pronouns include the pronouns *diaorang* 'they' and *luorang* 'you (PL)'. Bazaar Malay, Baba Malay, and Hokkien thus share the use of a noun – *orang* (Malay) or *lang* ~ *nang* / *-n* (Hokkien) – as pronominal plural marker. According to Paauw (2008: 170), the use of *orang* (or its variants) as plural marker is also observed in a number of eastern Indonesian contact varieties: Manado Malay (except in the 2PL), Ambon Malay, Banda Malay, North Moluccan Malay (except in the 2PL), Kupang Malay, Larantuka Malay, and Papua Malay (except in the 2PL).

Loan translations are also observed in possessive constructions. In generic Malay, possession is formed syntactically, by placing the respective noun or pronoun after the head noun (the possessed item). For example, *buku Nora* 'Nora's book' or *rumah saya* 'my house'. The third person pronouns *dia* 's/he' and *mereka* 'they' have a different form (*-nya*) to indicate possessives, e.g. *rumahnya* 'his/her/their house'. In Bazaar Malay and Baba Malay, conversely, possession is expressed through the Malay word *punya* 'to possess' after the respective pronouns or nouns. Consider the following examples from, respectively, Baba Malay taken from the second chapter of *Si Hitam Yang Chantek* – in which 782 instances of such use were observed – and Bazaar Malay:

4) *Saya*   *keluarkan*   *semua*   *saya*   *punya*   *kekuatan*
    1SG   exert   all   1SG   punya   strength
    'I exerted all my strength'

5) *Dia*   *kasi*   *dia*   *punya*   *kawan*   *hantar*.
    3SG   CAU   3SG   punya   friend   send
    'She let her friend send (it)'. (Bazaar Malay Informant G)

The following juxtaposition demonstrates clearly how this construction is copied from Hokkien, whereby the Hokkien particle *e* 的 is added immediately after the possessor noun or pronoun and replaced with the Malay word *punya* 'possess' in the two contact languages:

	PRON	**+ POSS**
Hokkien	*i* 伊 's/he'	*i e* 伊的 'his; her'
Bazaar Malay, Baba Malay	*dia* 's/he'	*dia punya* 'his; her'

The same phenomenon, that is, the use of the possessive marker *punya* or its variants, is observed in the contact varieties of eastern Indonesia (Paauw (2008: 174)): *pe* in Manado Malay and North Moluccan Malay, *pung* or *pong* in Ambon Malay, *pung* or *pu* in Banda Malay, *pung* in Kupang Malay, *puN* in Larantuka Malay, and *punya* or *pu* in Papua Malay. Together with the second person pronoun *lu* in Kupang Malay stated earlier, this suggests an eastward influence ultimately exerted by Hokkien.

Temporal expressions in Bazaar Malay and Baba Malay also demonstrate the relexification of Hokkien. To name the days of the week, both Baba Malay and Bazaar Malay use the combination of the noun *hari* or *ari* 'day' and a number. Although these designations consist of Malay words, they are arranged in a way that reflects a Hokkien pattern. The naming system for the days of the week in generic Malay, Bazaar Malay, Baba Malay, and Hokkien is illustrated in Table 4.2.

While Pakir (1986: 113) suggests that this naming system in Baba Malay is an outcome of Malay lexification, I see it as another instance of lexical calquing, with Hokkien entries being replaced by Malay parallels. The generic Malay naming system is itself derived from Arabic numerals, *Isnin* 'two', *Selasa* 'three', etc., but with a different numbering. For example, while Monday is referred to as 'Day 1' in Baba Malay and Bazaar Malay, it is referred to as 'Day 2' (*Hari Isnin*) in Standard Malay. While the use of numerals to refer to the weekdays

TABLE 4.2 Expressions for the days of the week in the languages concerned

Generic Malay	Bazaar Malay	Baba Malay	Hokkien	English gloss
Hari Isnin	(h)ari satu	(h)ari satu	pai-it 拜一	Monday
Hari Selasa	(h)ari dua	(h)ari dua	pai-zi 拜二	Tuesday
Hari Rabu	(h)ari tiga	(h)ari tiga	pai-sa 拜三	Wednesday
Hari Khamis	(h)ari empat	(h)ari empat	pai-si 拜四	Thursday
Hari Jumaat	(h)ari lima	(h)ari lima	pai-go 拜五	Friday
Hari Sabtu	(h)ari enam	(h)ari enam	pai-lak 拜六	Saturday
Hari Ahad / Minggu	(h)ari tujuh / minggu	(h)ari minggu	pai-chit 拜七 / lei-pai 禮拜	Sunday

is common cross-linguistically, the Sinitic system was adopted in Baba Malay and Bazaar Malay. Following the Hokkien morphosyntactic property whereby the noun *pai* 拜 'day' is followed by the number of the specific day in the week, Bazaar Malay and Baba Malay thus replaced these Hokkien words with their Malay counterparts: *hari ~ ari* 'day' followed by a number. As noted by Lee (2014), the names of the months in Baba Malay likewise follow the Hokkien structure replaced with the corresponding Malay lexical items, as in *bulan satu* 'January' (lit. 'month one').

Bazaar Malay, Baba Malay	*hari satu*	'Monday'
	day one	
Hokkien	*pai-it* 拜一	'Monday'
	day-one	

See below for an example from *Si Hitam Yang Chantek* (Chapter 43):

6) *Dan   pada hari minggu saya   tak   kerja,   dan   bila   hari satu*
   and   on   Sunday   1SG   not   work   and   when   Monday
   datang   saya   rasa   segar.
   come   1SG   feel   fresh
   'And on Sunday, I do not work, and when Monday comes I feel fresh.'

It is not uncommon to hear the same naming system used by ethnic Malays in informal contexts. More research is needed to investigate whether this is a feature of informal Malay, which tends to be semantically more transparent than standard Malay, or the result of influence from Bazaar Malay as a historical lingua franca.

5.3   *Loan Blends*

A loan blend is a hybridization of foreign and native elements. Examples of loan blends in English, as given by Tseng (2004: 170), include 'chopstick', where *chop* is a Chinese Pidgin English word; 'hoisin sauce', where *hoisin* 海鮮 is Chinese; and 'kung pao chicken', where *kung pao* 宮保 is a Chinese element. Though this type of loan blend was not observed in my more formal Bazaar Malay data, I have noticed its usage in informal conversations with older shopkeepers in the wet market.

In Baba Malay, out of the 72 kinship terms listed by Lee (2014: 399–402), we find ten loan blends which combine a Hokkien kinship term followed by a Malay noun or cardinal number indicating the seniority of the particular

relative. Examples include *tachi besar* 'eldest sister' (lit. elder.sister big) and *hia numbor dua* 'second elder brother' (lit. elder.brother number two), containing the Hokkien words *tachi* 大姐 and *hia* 兄. Some other expressions include *ati it gor it chap* 'nervous', where the Malay word *ati* 'heart' is blended with the Hokkien words *it gor it chap* 一五一十 'one five, one ten', and *buat suay* 'to cause misfortune' (lit. to make misfortune), where the Malay word *buat* 'to make' is blended with the Hokkien word *suay* 衰 'misfortune'.

6 Concluding Remarks

This chapter has highlighted two aspects relating to the role of Hokkien in the lexicon of Bazaar Malay and Baba Malay. The first is the extent of Hokkien influence on Bazaar Malay in comparison to Baba Malay, and the second the nature of Hokkien lexical influence in Bazaar Malay from the perspective of Pidgin and Creole Linguistics. Out of the four options of borrowing I have discussed, most instances involve the direct use of Hokkien words and loan translations, which can be observed in both varieties. Limited examples of loan shifts and loan blends occur in Baba Malay, but not in Bazaar Malay.

The influence of Hokkien on the lexicon of these two contact varieties is both overt and covert. Compared with Baba Malay, its lexical impact on Bazaar Malay is minimal. Bazaar Malay retains a small number of overt lexical items from Hokkien which can be categorized into four groups (compared to twelve groups attested in Baba Malay) that are clustered around cultural domains. Even in these four categories, there are fewer Hokkien words in Bazaar Malay than in Baba Malay.

This can be explained by two factors: a sense of community and of Baba Chinese identity. Baba Malay speakers belong to a distinct speech community and constitute a culturally identifiable group, historically different from China-born immigrants. Contrary to this, Bazaar Malay has never existed as a (native) speech community. Besides, as it is spoken by non-Malays of Chinese parentage but also other ethnic groups, we cannot expect the same degree of Hokkien lexical retention seen in Baba Malay. In the latter, lexical items closely tied with the Hokkien Chinese culture, customs, emotions, and ethnic value judgments have been handed down as cultural heritage by the ancestors of today's Baba Malay, although nowadays they do not typically speak Hokkien or any other regional Sinitic varieties anymore (see also Tan 1988: 120). These lexical items, which do not exist in Bazaar Malay, in fact mark the unique identity of the Baba Chinese as opposed to non-Baba Chinese. The survival of such

lexical items in their language can thus be accounted for by their persistence to maintain the Baba Chinese identity and their sense of belonging to the associated speech community.

In comparison with other pidgins and creole languages, the degree of influence from Hokkien in Bazaar Malay is remarkably significant. Since the speakers of substrate languages tend to be stigmatized in contact situations, as argued by Bao (2003), borrowing from the substrate is rare or minimal cross-linguistically. The presence of Hokkien words in Bazaar Malay confirms that Hokkien speakers were not historically stigmatized, given their social and economic dominance in the region. In this regard, the significant role of Hokkien in Bazaar Malay exemplifies an unusual scenario of language contact, in which the substrate language and its speakers receive a better treatment than is usually the case.

As a final point, we may recall that typical creoles are developed in a contact situation involving two groups with unequal power. The presence and nature of Hokkien loanwords in Baba Malay have led me to agree with Ansaldo et al. (2007) and Lee (2014: 22), namely, that Baba Malay was born out of intermarriage between ethnicities of relatively equal power. This linguistic finding has broader implications for historical scholarship on the Malay-speaking world and invites comparisons elsewhere in Southeast Asia.

References

Adelaar, K. Alexander, 'Some notes on the origin of Sri Lanka Malay', in H. Steinhauer (ed.) *Papers in Austronesian Linguistics* No. 1 (Canberra: Department of Linguistics, Research School of Pacific Studies, 1991): 23–37.

Adelaar, K. Alexander and Prentice, D.J., 'Malay: Its history, role and spread', in Stephen A. Wurm, Peter Mühlhäusler and Darell T. Tryon S. (eds.) *Atlas of languages of intercultural communication in Pacific, Asia and the Americas*, Volume II/1 (Trends in Linguistic Documentations 13, Berlin: Mouton de Gruyter, 1996): 6736–93.

Ansaldo, Umberto and Matthews, Stephen, 'The Minnan substrate and creolization in Baba Malay', *Journal of Chinese Linguistics* 27/1 (1999): 37–68.

Ansaldo, Umberto, Lisa Lim, and Salikoko S. Mufwene, 'The sociolinguistic history of the Peranakans: What it tells us about "creolization"', in U. Ansaldo, S. Matthews and L. Lim (eds.), *Deconstructing Creole* (Amsterdam: John Benjamins, 2007): 203–26.

Aye, Khin Khin, *Bazaar Malay: history, grammar and contact* (unpublished PhD thesis, National University of Singapore, 2006).

Aye, Khin Khin, 'Singapore Bazaar Malay', in Susanne Michaelis, Philippe Maurer, Martin Haspelmath, and Magnus Huber (eds.) *The survey of pidgin and creole*

languages Volume III: contact languages based on languages from Africa, Australia, and the Americas (Oxford: Oxford University Press, 2013): 86–93.

Bakker, Peter, 'Bazaar Malay & Baba Malay: Is There a Connection?', Paper presented at the *SPCL Conference at the 2004 LSA Meeting* (Boston, Massachusetts, 2004).

Bao Zhiming, 'Social stigma and grammatical autonomy in nonnative varieties of English', *Language in Society*, 32 (2003): 23–46.

Bao Zhiming, 'The aspectual system of Singapore English and the systemic substratist explanation', *Journal of Linguistics* 41/2 (2005): 237–67.

Bao, Z.M., 'The systemic nature of substratum transfer', in Markku Filppula, Juhani Klemola and Devyani Sharma (eds.), *The Oxford Handbook of World Englishes* (Oxford: Oxford University Press, 2015): 621–38.

Baxter, A.N., *A description of Papia Kristang* (*Malacca Creole Portuguese*). (PhD thesis, Australian National University, Canberra, 1985).

Bodman, Nicholas C., *Spoken Amoy Hokkien* (2 vols, Kuala Lumpur: Charles Grenier and Company, 1955).

Collins, J.T., 'Malaysian and Bazaar Malay: polarity, continuity and communication', in Asmah Haji Omar (ed.), *National language and communication in multilingual societies* (Kuala Lumpur: Dewan Bahasa dan Pustaka, 1987): 151–74.

Collins, J.T., *Malay, world language: A short history* (2nd edn., Kuala Lumpur: Dewan Bahasa dan Pustaka, 1998).

Durkin, Phillip, *Borrowed words: A history of loanwords in English* (Oxford: Oxford University Press, 2014).

Egerod, S., 'A Note on Some Chinese Numerals as Loan Words in Tai'. *Second Series*, 47/1–2 (1959): 67–74.

Freedman, Maurice, 'Immigrants and associations: Chinese in the 19th Century Singapore', *Comparative Studies in Society and History* 3 (1960): 25–48.

Gil, David, 'Escaping Eurocentrism: Fieldwork as a process of unlearning', in Paul Newman and Martha Ratliff (eds.), *Linguistic fieldwork* (Cambridge: Cambridge University Press, 2001): 102–32.

Gwee, Willian Thian Hock, *Mas sepuloh: Baba conversational gems* (Singapore: Armour Press, 1993).

Gwee, Willian Thian Hock, *A Baba Malay dictionary: The first comprehensive compendium of Straits Chinese terms and expressions* (Singapore: Armour Press, 2006).

Hall, K.R., 'The opening of the Malay world to European trade in the sixteenth century', *Journal of the Malaysian Branch of the Royal Asiatic Society* (1985): 85–104.

Hay, J., 'Chinese in Southeast Asia', *Facts and Details*, 2008. Retrieved 30 June 2008 from http://factsanddetails.com/asian/cat66/sub418/item2729.html.

Hockett, C.F. (ed.), *The life and growth of language* (New York: Dover Publication, 1979).

Hoffer, B.L., 'Language borrowing and the indices of adaptability and receptivity', *Intercultural Communication Studies*, XIV/2 (2005): 53–72.

Holm, John, *Pidgins and creoles*, vol. 1 (Cambridge: Cambridge University Press, 1989).

Hoogervorst, Tom G., 'What kind of language was "Chinese Malay" in late colonial Java?', *Indonesia and the Malay World*, 45/133 (2016): 294–314.

Jones, Russel, *Chinese Loan-Words in Malay and Indonesian: A Background Study* (Kuala Lumpur: University of Malaya, 2009).

Kuo, Eddie Chen-Yu, 'A sociolinguistic profile of Singapore', *Sociology Working Paper* No. 52 (Singapore: Chopmen Enterprises, 1976).

Lee, Nala Huiying, *A grammar of Baba Malay with Sociophonetic considerations* (unpublished PhD thesis, The University of Hawaii, 2014).

Lefebvre, Claire, *Creole genesis and the acquisition of grammar: The case of Haitian creole* (Cambridge: Cambridge University Press, 1998).

Lefebvre, Claire and Lumsden, John, 'Le role central de la relexification dans la genèse de langues creoles', *Plurilinguismes* 8 (1994a): 47–93.

Lefebvre, Claire and Lumsden, John, 'Relexification in creole genesis', in Claire Lefebvre and John Lumsden (eds.), *The central role of relexification in creole genesis: The case of Haitian creole* (research report prepared for SSHRCC on the project *La genèse du créole haïtien: un cas particulier d'investigation sur la forme de la grammaire universelle*, Université du Québec à Montréal, 1994b).

Leow, Rachel, *Taming Babel: Language in the Making of Malaysia* (Cambridge: Cambridge University Press, 2016).

Ling Pyau, 'Causes of Chinese Emigration', *The Annals of the American Academy of Political and Social Science*, 39 (1912): 74–82.

Lockard, Craig A., 'Chinese migration and settlement in Southeast Asia before 1850: Making fields from the sea', *History Compass* 11/9 (2013): 765–81.

Mandal, Sumit, 'Cultural Geographies of the Malay World: Textual Trajectories in the Indian Ocean', *Philological Encounters*, 1 (2016): 370–95.

Milroy, Lesley and Gordon, Matthew, *Sociolinguistics: Method and interpretation* (Malden, MA: Blackwell, 2003).

Minde, D. van, *Malayu Ambon: Phonology, morphology, syntax* (Leiden: Research School CNWS, 1997).

Mithun, Marcyliena, 'Who shapes the record: The speaker and the linguist?' in Paul Newman and Martha Ratliff (eds.), *Linguistic fieldwork* (Cambridge: Cambridge University Press, 2001): 34–44.

Muysken, Pieter, 'Half-way between Quenchua and Spanish: The case for relexification', in Arthur R. Highfield and Albert Valdman (eds.), *Historicity and variation in Creole Studies* (Ann Arbor, MI: Karoma, 1981): 52–79.

Ng, Kim Chew, 'Sinophone/Chinese: "The South where Language is lost" and Reinvented, translated by Brian Bernards', in Shu-mei Shih, Chien-hsin Tsai, and Brian Bernards (eds.) *Sinophone Studies: A Critical Reader* (New York: Columbia University Press, 2013), 74–92.

Paauw, Scott, *The Malay contact varieties of eastern Indonesia* (unpublished PhD thesis, State University of New York, 2008).

Pakir, Anne, 'Lexification and idiomaticity in Baba Malay, the language of the Straits Chinese', *Working Papers in Linguistics* (Honolulu: University of Hawaii, Department of Linguistics), 16/2 (1984): 63–85.

Pakir, Anne, *A linguistic investigation of Baba Malay* (unpublished PhD thesis, University of Hawaii, 1986).

Pan, Lynn, *Sons of the Yellow Emperor: A history of the Chinese diaspora* (Boston, MA: Little, Brown, 1990).

Platt, John and Weber, Heidi, *English in Singapore and Malaysia: Status, features, functions* (Oxford: Oxford University Press, 1980).

Poedjosoedarmo, Gloria R., 'What is happening to Malay in Singapore?', in Cecilia Odé and Wim Stokhof (eds.), *Proceedings of the Seventh International Conference on Austronesian Linguistics* (Amsterdam: Rodopi, 1997): 73–84.

Population Trends, 'Department of Statistics, Ministry of Trade and Industry, Republic of Singapore', 2017.

Poston, D.L. Jr., 'Overseas Chinese', in Demeny P. and McNicoll G. (eds.) *Encyclopedia of Population* (New York, NY: Macmillan Reference USA, 2003).

Poston, D.L. Jr. and Luo, H., 'Chinese student and labour migration to the United States: Trends and policies since the 1980s', *Asian and Pacific Migration Journal*, 16/3 (2007): 323–55.

Poston, D.L. Jr. and Wong, Juyin Helen, 'The Chinese diaspora: The current distribution of the overseas Chinese population', *Chinese Journal of Sociology*, 2/3 (2016): 348–73.

Pou, Saveros and Jenner, Philip N., 'Some Chinese loanwords in Khmer', *Journal of Oriental studies* (1973): 1–90.

Prentice, D.J., 'The best chosen language', *Hemisphere*, 22/3 (1978): 18–23.

Prentice, D.J., 'Malay and Indonesia', in William Bright (ed.), *International encyclopedia of linguistics* (Oxford: Oxford University Press, 1992): 374–80.

Proudfoot, I., *Early Malay printed books: A provisional account of materials published in the Singapore-Malaysia up to 1920, noting holdings in major public collections* (Kuala Lumpur: University of Malaya, 1993).

Purcell, Victor, *The Chinese in Malaya* (Oxford: Oxford University Press, 1948).

Sebba, Mark, *Contact languages: pidgin and creole languages* (New York: St. Martin's Press, 1997).

Shellabear, W.G., 'Christian Literature for Malaysia: I. Races and Languages', *The Moslem World*, IX/4 (1913): 379–84.

Tadmor, Uri, 'Loanwords in the world language: findings and results', in Martin Haspelmath and Uri Tadmor (eds.), *Loanwords in the world languages: A comparative handbook* (De Gruyter Mouton: Berlin, 2009): 36–55.

Tan Chee Beng, *The Baba of Melaka: culture and identity of a Chinese Peranakan community in Malaysia* (Selangor: Pelanduk Publications, 1988).

Tan Chee-Beng, *Chinese overseas: comparative cultural issues* (Hong Kong: Hong Kong University Press, 2004).

Tarling, Nicholas, *Singapore and the Singaporeans since 1819* (Auckland: University of Auckland, 1992).

Thurgood, Elzbieta A., *A description of Nineteenth Century Baba Malay: A Malay variety influenced by language shift* (PhD thesis, University of Hawaii. Ann Arbor, 1998).

Tseng, Hugo T.Y., 'Rendering Chinese-culture specific vocabulary into English – Predicaments and prospects', in Chan Sin-Wai (ed.), *Translation and Bilingual Dictionaries* (Tübingen: Niemeyer Max Verlag GmbH, 2004).

Vaughan, Jonas Daniel, *Manners and customs of the Chinese of the Straits Settlements* (Kuala Lumpur: Oxford University Press, 1971).

Wang, Anyi, 'The Fate of Language (Parts 1 and 2)', *Xingzhou ribao, xingyun ban* [Sin Chew daily starcloud supplement] (13 July and 7 December 1993).

Wang, Gungwu, *China and the Chinese overseas* (Singapore: Times Academic Press, 1991).

Yoong, S.K. and Zainab, A.N., 'Chinese literary works translated into Baba Malay: A bibliometric study', *Malaysian Journal od Library & Information Science*, 7/2 (2002): 1–23.

Native Lexical Innovation in Penang Hokkien: Thinking beyond *Rojak*

Catherine Churchman

1 Introduction[1]

A few years ago, in a box of yellowed papers at the back of a cluttered antique shop in Chulia Street in George Town, I came upon a small green volume entitled "Chinese New Terms and Expressions" that had been published in Shanghai in 1913. The author, Evan Morgan, had spent several years collecting and noting down recently coined Chinese words he came across in newspapers, magazines, and books, and his work was a testimony to the rapid changes that had occurred in the Chinese written and spoken language over the previous decades, as China transformed itself from empire to nation-state and words for new technology and new ideas had entered the language. Having spent the previous few weeks wandering around George Town collecting vocabulary for a dictionary of the Hokkien language as spoken in Penang,[2] as I leafed through this volume it occurred to me that although many of the new Chinese terms it recorded were shared with other varieties of Hokkien – such as those spoken in Amoy and Taiwan – a fair number of these had not gained currency in the Hokkien of Penang. As Amoy and Taiwanese varieties have tended to follow the lead of Japanese and Mandarin in the creation of their modern vocabularies, Penang Hokkien vocabulary has, to some extent, modernized along a different trajectory. This is due in part to Penang Hokkien speakers' longstanding acceptance of loanwords from Malay and English – hence the common metaphor *rojak* 'spicy fruit-and-vegetable salad' – and their free use of these in place of native Hokkien vocabulary, but another significant contributing

---

1 All Hokkien words are transcribed according to the POJ or Church Romanization system with significant modifications for Penang pronunciation.

2 Although it is not used in either China or Taiwan, I employ "Hokkien" in its Southeast Asian sense as a catch-all for the spoken idioms of the Zhangzhou-Quanzhou districts of southern Fujian. The terms "Southern Min" or "Minnan" 閩南 as used in China can have a wider meaning encompassing related languages that not only includes the Zhangzhou-Quanzhou group of languages, but also Teochew and Hainanese (Zhou 2010: 1).

factor to the divergence is Penangites' creative compounding of pre-existing Hokkien vocabulary in ways unknown to speakers of other varieties. The following describes both the historical context of this divergence in vocabulary, and examples of Penang Hokkien speakers' neologisms arranged according to some general categories.

Despite the name, "Penang Hokkien" as employed in this chapter refers not only to varieties of Hokkien spoken within the state of Penang, but also to closely related varieties spoken by the majority of Malaysian ethnic Chinese in the cities and towns along the west coast of the Malay Peninsula, from the north of Kuala Kangsar to the Thai border. Because of its geographical spread, Penang Hokkien is also sometimes known as Northern Malaysian Hokkien. This serves to distinguish it from the "Southern Malaysian Hokkien" spoken in the southern half of the Malay Peninsula and in Singapore, which is phonologically and grammatically different from the Northern variety.

Yet other more distantly-related varieties can be found further afield, in Northern Borneo (Sarawak), Northern Sumatra (Medan) and Southern Thailand (Phuket), and a related variety was also once spoken in the Chinatown of Rangoon. However, within Malaysia, Penang remains the most important centre for this language and was historically the centre from which it spread to other areas along the coast. Penang Hokkien has been the lingua franca of ethnic Chinese in northern Malaysia for over a century, but at present appears to be undergoing replacement by Mandarin as a result of longstanding education policies and changes in the home linguistic environment (Sim 2012; Ooi & Tan 2017).

Penang Hokkien is based on the Haicheng 海澄 dialect of the Zhang-Quan 漳泉 subgroup of Southern Min languages, reflecting the speech of the homeland of the majority of the earliest Hokkien settler families prior to the middle of the nineteenth century (Jones 2009). However, intense contact with other languages, principally Malay, English, Cantonese, and Teochew, has significantly altered the vocabulary, tonal system, and syntactic structure of the base dialect to a greater of lesser degree that varies in accordance with a speaker's social network, age, and family and educational backgrounds (see Lim & Teoh 2007; Lim 2010; Chuang et al. 2013; Soon 2014; Ye 2014; Churchman 2017; Hing 2017). The divergent development of all of these linguistic features in Penang Hokkien frequently impedes mutual intelligibility with Taiwanese and mainland Chinese varieties.

As the most immediately obvious difference between Penang Hokkien and other varieties, it is lexical borrowings from Malay that attracts the most scholarly and popular attention. Social media discourse around the language often revolves around such words, the perennial favourites being *lui* 鐳 for 'money'

and *sah-bûn* 雪文 for 'soap' and their supposed derivations.[3] However, lexical borrowing is only one of the factors behind the distinctive character of Penang Hokkien vocabulary, as a substantial amount of vocabulary difference between Penang and other varieties originates in different usages of native Hokkien vocabulary. One category of divergent usages is the retention in Penang of a number of words that have fallen out of use in other varieties. Two common examples are *chít-chûi* 一誰 for 'who' – used in the Hokkien translation of the Old Testament completed in 1884 – and *thiⁿ-sî* 天時 for 'weather', both current in Penang but already obsolete in Taiwan and Amoy usage (the latter is known in Singapore). A further factor, much understudied and the subject of this chapter,[4] is the rich vocabulary of neologisms that speakers of Penang Hokkien have derived through compounding and semantic extension of the elements shared with other varieties of Hokkien. These words, although derived from native Hokkien elements, are mostly unknown in the Hokkien-speaking world beyond Malaysia and Singapore and are generally not understood by Taiwanese or Amoy speakers of Hokkien.

Such neologisms can be categorized into two broad categories. The first includes words for things that speakers encountered in the multicultural, tropical environment of the Straits Settlements that were rarely encountered by or unknown to speakers of Hokkien in nineteenth-century Fujian, including types of clothing, building styles, foods, fruits, religious practices, customs, and colonial political or administrative institutions. The second category consists of words for the new technology and ideas that appeared in industrialized societies of the twentieth century. These were commonly encountered in all parts of the Hokkien-speaking world, but were invented or came into existence at a time in which the Penang Hokkien speech community was relatively isolated from other Hokkien speech communities. Over the last ten years I have been collecting these distinctive terms from conversations on the street, internet forums, the weekly Penang Hokkien Podcast, and a collection of published dictionaries and manuals of the language (Tan 2001; 2008; 2010; Kwok 2005; Lee Eng Kew 2007; de Gijzel 2013; Tan: 2016). I am also extremely grateful for the help of Sim Lee, Ooi Kee How, Simon Chee Hooi Lim, Khoo Salma, Tina Teoh and Alan Ong for making me aware of many of these special terms.

---

3   The former comes from Dutch *duit* via Malay and the latter either from Arabic or Portuguese *sabun*. These and subsequent Malay etymologies are from Jones (2007).

4   Wu (2014) appears to be the only pre-existing academic study of such words in Penang. Gao (2000) has made a list of unique native terms in the closely-related dialect of Medan and Zhou and Zhou (2000: 89–90) provide a short list of neologisms in Singapore Hokkien.

2   The Development of New Vocabulary in Chinese Languages

The last two decades of the nineteenth century and the first decade of the twentieth were an important time for the development of Modern Standard Chinese vocabulary. As Chinese began to take an interest in the technology, culture, and governmental systems of Western countries, hundreds of new words were created and imported into written Chinese, and from there made their way into the spoken vernacular languages of those literate in Chinese (including Vietnamese and Korean). Beginning in the 1840s, Chinese intellectuals created new Chinese compound words they came across in Western books, and after the 1860s principally on the models provided by Japanese *kanji* compounds coined by Japanese writers, as Meiji Period Japan (1868–1912) quickly became a model for China of a successfully modernized, strong Asian nation (see Masini 1993). New terms in Chinese included words such as *minzhu* 'democracy' (appearing first in Chinese in 1864), *zongjiao* 'religion' (1890), *kexue* 'science' (1896), and *dianbao* (1860s) 'telegraphy', based on Japanese *minshu* 民主, *shūkyō* 宗教, *kagaku* 科學, and *denpō* 電報. These words eventually found their way into the spoken Chinese vernacular languages, and formed an important part of the vocabulary of the National Language (*Guoyu* 國語) based on the Peking dialect, promoted from the 1920s onward. Many words invented or borrowed into written Chinese over this period found their way into spoken Hokkien through the media and the education system. In the preface to his 1923 supplement to Carstairs Douglas' Hokkien dictionary of 1873, Thomas Barclay describes the changes in Hokkien vocabulary of the preceding half-century as follows:

> Western civilization, to an increasing extent, has been welcomed, and new ideas in every department of thought and action have filled the minds of the people. These new ideas have demanded for their expression new terms, the addition of which has much enriched and extended the language.
>
>   BARCLAY 1923: i

Barclay's supplement added 271 pages of entries to Douglas' original 617 pages, an increase somewhere between a third to a half of new vocabulary. Hokkien speakers on Taiwan during the Japanese colonial period (1895–1945) were exposed directly to newly coined Japanese vocabulary through the Japanese education system, and borrowed words like *iá-kiû* 野球 'baseball', *kháu-chō*

口座 'bank account', and *pēn-īn* 病院 'hospital' – that were modelled directly on Japanese *yakyū*, *kōza*, and *byōin* – entered the Taiwanese Hokkien vernacular but were not adopted in written Chinese. After 1945, the exclusion of languages other than Mandarin from the education system in Taiwan intensified the influence of Mandarin on Taiwanese Hokkien, with the result that many vernacular Taiwanese expressions were replaced by their Mandarin equivalents pronounced according to Hokkien, such as the replacement of the native Taiwanese *chò-sit-lâng* 做穡儂 'farmer' with the Mandarinized term *lông-bîn* 農民 (Cheng 1987). Similar changes occurred in the Hokkien spoken in China. Li and Xu (2007) analyse changes in Amoy dialect vocabulary over a century as illustrated through Amoy dialect textbooks, noting not only changes in material and social culture as reasons for vocabulary change, but also the strong influence of Mandarin vocabulary, especially from the mid-twentieth century onwards.

During this period of rapid change in the vocabulary of Chinese languages, three factors combined to encourage Penang Hokkien to develop divergently from other varieties. The first and most basic factor was the disconnect between spoken Hokkien and forms of written Chinese, exacerbated in Penang by the heavy admixture of Malay, English, and other languages. The second was historically low levels of literacy amongst the Penang Hokkien speech community in written Chinese or proficiency in Mandarin, the two media through which neologisms entered other varieties of Hokkien. The third factor was the almost complete lack of exposure to other varieties of Hokkien and the relegation of Penang Hokkien to the home, the family, and everyday societal interactions. Until the late 1980s, little was published in Hokkien worldwide aside from a few textbooks and missionary texts, let alone in Penang Hokkien, nor was there much in the way of news media throughout most of the twentieth century to carry new vocabulary current in Chinese or Taiwanese Hokkien to the ears of Penang Hokkien speakers. These three factors are explained in more detail below.

Until relatively recent attempts at standardization by the ROC Ministry of Education, spoken varieties of Hokkien have historically maintained only a loose relationship with the Chinese written character. In pre-twentieth century texts, character usage was largely unsystematized (see Klöter 2005: 41–87 for details) and the Literary Chinese used in the education system and as the official written language of the ROC until 1919 presented extra problems for Hokkien speakers on account of its pronunciation being significantly different from what they used in their everyday vernacular. Writing about Chinese

education in Malacca in 1839, Newbold quotes a report on a Chinese native school in Malacca describing the difficulties faced by those learning to read and write through Hokkien as follows:

> In schools among the Fokien people, the practice of committing much to memory is not attended with so much benefit as might be expected, from the circumstance of their colloquial dialect being entirely different from that in which they read and learn; insomuch, that though persons may be well acquainted with the colloquial dialect, yet the dialect in which they read is so different, that much may be committed to memory without being understood. This forms a great barrier to improvement in Fokien schools, as the scholars have two dialects to acquire before they can understand or make themselves intelligible to others. The same is the case in Canton schools.
>
> NEWBOLD 1839, vol. 1: 177

The "two dialects" in this case most likely refers to the difference between Literary Chinese and the colloquial vernacular. Both vernacular Cantonese and Hokkien differ greatly in grammar and vocabulary from Literary Chinese. However, for Cantonese speakers the reading pronunciation of texts and the pronunciation of the related morphemes in the spoken vernacular is largely identical, and it is only the word choice and grammatical idiom that distinguishes it from the literary style (see also Ding, this volume). For a Cantonese speaker, 'meat', 'person', and 'water' are pronounced *yuk* 肉, *yan* 人, and *sui* 水, no matter whether they are spoken or read off the page of a Chinese book. In contrast, a Hokkien speaker will use *bah*, *lâng*, and *chúi* when speaking, but when reading out from a page of Literary Chinese, read the Chinese characters out as *jiók*, *jîn*, and *súi*. It has been estimated that 33% of commonly used characters for written Taiwanese (that is, the written representation of the spoken vernacular) have a double reading, or more accurately, one character stands for what would be two separate morphemes in a phonetic script (Xu 2000: 62). This added an extra layer of complexity to an already complicated and ineffective language learning process. The result was that the literacy rate in Fujian was especially low. Edwin Joshua Dukes, a missionary who lived in southern Fujian in the 1880s describes it the lowest in the country:

> In the northerly provinces, where the Mandarin language is read and spoken, the proportion of readers is higher than in the south, where the dialects and languages are so numerous, and the written language coincides so little with the market tongues. The most deplorably ignorant

province is Fu-kien. Intelligent and judicious colporteurs in that province have assured the writer that only one or two percent can read with sufficient hope that if they received the scriptures their own eyes could convey the meaning to their eyes and hearts.

dukes 1885: 166–67

Newbold noted that in Malacca, Chinese tended to intermarry with Malays, and that the Malay language became the home language and the one acquired by children "to whom the later acquirement of Chinese must become a matter of time and difficulty" (Newbold 1839, vol. 1: 172). Later he notes that although Chinese from China were generally literate, "of those born in Malacca … probably not one in ten (though they have been at school several years) is able to understand books written in the plainest style" (ibid. 179–80).

Learning to read Chinese, already an onerous task for a Hokkien speaker, was further compounded in Penang by the fact that many Malay terms and a certain number of Malay syntactic features had entered the language, especially that of the Baba families (Lim & Teoh 2007, Lim 2010). In the last decades of the nineteenth century and the first decades of the twentieth, many Straits-born Chinese in Penang spoke Malay as their mother tongue rather than Hokkien. In 1913, William Shellabear refers to Baba Malay (a creolized form of Malay with a Hokkien substrate; see Aye in this volume) as "the mother-tongue of the majority of the Chinese women and children in the Straits Settlements", and "the language of the homes of the Straits-born Chinese – the most highly educated and most influential section of the Chinese community in the British possessions" (Shellabear 1913: 52). Judging by the popularity of the Malay translations of "Batu Gantong" – the pen-name of Chan Kim Boon 曾錦文 (1851–1920) – Chinese who could speak Hokkien were more comfortable reading and writing in Malay or English than any form of written Chinese, and wrote fan letters to him in both of these languages. One of them, named Lim Tiouw Chuan from Taiping, recommended his *Sam Kok* 三國 as a work "that should be in every Chinese home" (Batu Gantong 1892, vol. 5: iii–iv). Unlike many of his readers, Chan himself was a proficient reader of Chinese who had spent some years in Fujian, and his works contained glossaries of difficult or obscure Hokkien vocabulary explained in English and vernacular Malay.

Further complexity was added as a wave of migration of Chinese from China (known as the *Sinkhek* 新客; lit. 'new guests'), beginning in the second half of the nineteenth century, brought Chinese speaking other languages such as Cantonese and Teochew who introduced many features of their own native tongues into the Penang Hokkien. Throughout this period, English was the language of colonial administration and higher education, which resulted in

English vocabulary and syntactic features entering the language, particularly from the mid-twentieth century onwards (Churchman 2017). Amidst these diverse origins of the Penang Hokkien vernacular, even a Penang Hokkien speaker who was a proficient reader and writer of Classical Chinese or Mandarin would be at a loss how to turn their everyday Hokkien vernacular to a written form in Chinese characters. Consequently, until the rise of interest in vernacular Penang Hokkien beginning in the early 2000s, written representations of the spoken vernacular were rare and limited to a few wordlists, mentions and descriptions. Two notable examples are Lo Man Yuk's *Chinese Names of Streets in Penang* that provides a snapshot of vernacular vocabulary at the turn of the twentieth century (Lo 1900), and the Hokkien-Malay glosses contained in the Baba Malay translations of Chinese classic novels by Batu Gantong over three decades of the 1890s to the 1910s.[5] Both of these writers used romanized forms of Hokkien accompanied by what they considered to be the appropriate Chinese characters.

Hou Hongjian's 侯鴻鑒 *A Record of Travel in the Nanyang*, published in 1920, offers a snapshot of Chinese language education in the first two decades of the twentieth century. Hou recounted his discussions with the headmasters of Chinese schools in Penang in October 1919, in which he was informed that there were twenty Chinese primary schools in Penang, fifteen boys' schools and five girls' schools, but that only five of these had rolls of more than 100 pupils, the Chunghwa Confucian School 孔聖廟中華國民型中學 boasting the highest roll of over 400 pupils (Hou 1920, vol. 2: 30a). Hou noted that in the Chunghwa School taught the first two years in Hokkien before switching to the "National Language" (Mandarin) in the third and fourth years (ibid., vol. 2: 31b.). Founded by the local business magnate Cheong Fatt Tze 張弼士 in 1904, this school was the first to use Mandarin as the medium of education but was likely teaching literacy in Classical Chinese rather than Vernacular Chinese (白話 *Baihua*) – a written form based on northern Chinese – as the written standard.

The Republican government in China began to popularize *Baihua* as the standard written form of Chinese from the 1920s onwards, and this began to filter into the Chinese education system in Penang (also see Hoogervorst, this volume, on Indonesia). Although it is not clear when various Chinese-medium schools in Penang shifted to the new form, the Chinese newspaper *Penang Sin Poe* 檳城新報, published from 1895 to 1941, gradually switched from Classical Chinese to Mandarin over the period from 1928 to 1930. This at least indicates

---

5   Two of the most notable translations being his *Sam Kok* 三國 (Luo Guanzhong's 羅貫中 *Romance of the Three Kingdoms* published from 1892–96) and *Kow Chey Thian* 猴齊天 (Wu Cheng'en's 吳承恩 *The Journey to the West* published from 1911–13).

that the new style of reading and writing had become widespread by this time among those literate in Chinese. Although the adoption of *Baihua* reduced the distance between written Chinese and all of the spoken Sinitic languages to some extent, the distance between the structure and vocabulary of written Mandarin and that of spoken Penang Hokkien would still have constituted a considerable hurdle to Hokkien speakers' ability to relate what they had learnt to write in Mandarin back to their everyday language.

The total Chinese population of Penang in the year following Hou's visit was 135,288, and that of Hokkiens 64,085. Although it is impossible to calculate the number of proficient Chinese readers and writers in Penang, Hou's estimation from 1919 indicates that at the most, around 2,000 children were attending schools in which they could acquire any kind of Chinese literacy, and that this was a relatively new phenomenon. In Penang of the pre-War period, there were many people who had had no chance to attend schooling of any kind, and those who did often attended only for a few years or attended English-medium schools. In addition, Penang Hokkien speakers often had to use Malay or English for social interactions and for official purposes and in higher education. They read Malay and English newspapers and books.

Popular entertainment in other varieties of Hokkien could have acted as a conduit in introducing new words from these varieties to Penang speakers, but such entertainment was not widely available to them. Entertainment in Cantonese and Mandarin was easily available in Malaysia from the 1980s onwards, but Penang Hokkien speakers had relatively little exposure to entertainment and news in any variety of Hokkien until as late as the mid-2000s. Ten-minute nightly news broadcasts in Hokkien were and continue to be broadcast nationally, but these have always been made in the Amoy dialect that is far removed from what is spoken in Penang. This changed in 2007 with Taiwanese programming through satellite TV, such as Hua Hee Dai 歡喜台 started in 2007, or through sites such as Youtube.[6]

Therefore, during the period in which the vocabulary of written Chinese – and hence, other varieties of Hokkien – was in a state of flux, the disconnect between written Chinese and spoken Hokkien, the limited spread of education in written Chinese, and intense contact with English and Malay affecting mutual intelligibility with other varieties all combined to insulate the majority of Penang Hokkien speakers from the major developments in Hokkien vocabulary in other parts of the Hokkien-speaking world well into the twenty-first century. The result of these combined factors was that speakers of Penang

---

6   For more on the recent sociolinguistic situation of Chinese languages in Malaysia, see Sim (2012).

Hokkien were insulated from the trend of adopting words into Hokkien via graphic loans from Mandarin and instead derived original neologisms from their own everyday spoken language through compounding and semantic extension of pre-existing words. The vocabulary of Penang Hokkien therefore developed along a different trajectory from other varieties.

3   Vocabulary Innovation by Compounding

As Penang Hokkien compounds are derived from putting together common pre-existing words, they are more transparent in meaning to speakers than those borrowed from written Chinese in other varieties of Hokkien. Often in Literary Chinese, a single syllable will represent an idea but is only used as a bound morpheme, and unless one is familiar with its meaning and how it is used in other words, its meaning will be opaque. For instance, some of the examples in Table 5.1 relating to uniforms use the common morpheme *saⁿ* 衫 'shirt' for clothing in preference to *hók* 服 used in other varieties. As *hók* 服 appears only in the colloquial Penang Hokkien word *hók-sāi* 服侍, meaning 'to serve' or 'to worship', and not in words related to clothing, it did not occur to Hokkien-speaking Penangites when they came to derive words for uniforms.

Where Mandarin (and consequently Amoy Hokkien and Taiwanese due to their exposure to Mandarin) use two or three-character compounds to create concrete nouns, Penang Hokkien, lacking some of the syllabic morphemes transmitted through written Chinese, often prefers multisyllabic transparent compounds in which the sense of the word is immediately deducible from its constituent parts. Some terms are derived with the nominalizing particle *ê* 个 and are either short noun phrases containing subordinate clauses or possessive noun phrases (see Table 5.1). When referring to occupations, Penang Hokkien can derive a noun phrase from a verb object phrase with the addition of – *ê* "one who …". However, these are often avoided in preference to short sentences expressing the same concept, for example "I am a civil servant" would be expressed as *Wá chò chèng-hú-kang* 我做政府工, literally meaning 'I do government work'.

A number of very basic words are expressed differently in Penang Hokkien from other varieties of Hokkien. These single Penang words often cover a broader range of meanings than those in other varieties, which distinguish them with distinct terms. When these words are used as components in compound words, the meanings are often unclear to speakers of other varieties (Table 5.2).

TABLE 5.1   Transparent multisyllabic compounds

English meaning	Penang	Translation of Penang expression	Amoy/Taiwan[a]	Mandarin
army uniform	*peng--ê-sa^n* 兵个衫	'the clothes of soldiers'	*kun-hók* 軍服	*junfu* 軍服
chopstick holder	*hē·-tū--ê* (*míh-kià^n*) 下箸个(物件)	'the thing to put chopsticks in'	*tī-lāng* 箸籠	*kuaizitong* 筷子筒
civil servant	*chò chèng-hú-kang--ê* 做政府工个	'someone who works for the government'	*kong-bū-oân* 公務員	*gongwuyuan* 公務員
dentist	*pó·-chhùi-khí--ê* 補喙齒个	'someone who does fillings'	*gê-i* 牙醫	*yayi* 牙醫
fire extinguisher	*kiù-hóe--ê míh-kià^n* 救火个物件	'the thing that rescues from fires'	*biát-hóe-khì* 滅火器	*miehuoqi* 滅火器
ritual implements	*pài Hút--ê míh-kià^n* 拜佛个物件	'things for worshipping the Buddha'	*hoat-khì* 法器	*faqi* 法器
school uniform	*óh-tn̂g-(ê)-sa^n* 學堂衫	'the clothes for school'	*hāu-hók* 校服	*xiaofu* 校服
veterinarian	*khîm-siù--ê ló-kun* 禽獸个老君[b]	'doctor for animals'	*siū-i* 獸醫	*shouyi* 獸醫

a   The examples for Amoy and Taiwan in this table and those following have been cross-checked through the *Maryknoll English-Taiwanese Dictionary* (Maryknoll 2013), the *Minnan fangyan dacidian* (Zhou 2006), and the *Minnanhua Zhangqiang cidian* (Chen 2006).

b   *Khîm-siù* 禽獸 for 'animal' is obsolete in most other varieties or restricted to the meaning 'birds and beasts', having been replaced by the Japanese coinage *tōng-bút* 動物. *Ló-kun* 老君, the default term for 'doctor', is in fact the Malay loan *dukun*. This old term for a medicine man or medicine woman is disguised with the Chinese characters for 'elderly lord', the name of a Taoist deity.

TABLE 5.2   Differences in base morphemes

English	Penang Hokkien	Taiwan/Amoy	Meaning of Penang term in other varieties
certificate	*jī* 字	*chèng-bêng-su* 證明書, *jī* 字, *jī-bú* 字母	Chinese character, written word
deity, idol, divine	*ang-kong* 尪公	*sîn* 神, *sîn-bêng* 神明, *sîn-siōng* 神像	idol, doll
film, movie, TV programme	*hì* 戲[a]	*tiān-iáⁿ* 電影, *chiat-bók* 節目	a play
food, cuisine	*chiáh* 食	chhài 菜, liāu-lí 料理, chhan 餐	to eat, consume
temple, building for religious worship, monastery	*am* 庵	*biō* 廟, *sī* 寺,[b] *am* 庵	Buddhist monastery, nunnery
tree, plant	*châng* 欉	*chhiū-á* 樹仔, *chhiū-châng* 樹欉, *sít-bút* 植物	measure word for plants

a   *Tiān-iáⁿ-hì* 電影戲 is another older way of referring to a film, but the shorter form is more common.
b   Both of these terms are included in the names of temples in Penang, for example Kék-Lók-Sī 極樂寺, the famous Buddhist temple in Air Itam, and Chôa-Biō 蛇廟, the Snake Temple at Bayan Lepas. The word *biō* 廟 on its own is not unknown, however a Penang Hokkien speaker would tend to describe these collectively as *am* 庵.

The element *am* 庵 has the extended meaning of any large building used for religious worship, with the exception of Christian churches. Its derived compounds include *Hoan-á-am* 番仔庵 'a Mosque' (lit. Malay-*am*) – also *hôe-kàu-tn̂g* 回教堂 or *hôe-kàu-biō* 回教廟 – *Siām-Hút-am* 暹佛庵 'Siam Buddha-temple' for a Thai wat,[7] *Bâng-gà-lí am* 芒加里庵[8] (lit. 'Bengali temple') for a Sikh temple, and *Kè-lêng-á-am* 吉寧仔庵 for a Hindu temple.

The word *ang-kong* 尪公 referred to a doll or idol, or the statue or sculpture of a deity, and these are the principal meanings it retains in other varieties of Hokkien. However, in Penang and northern Malaysia, although the original

7   Even though Thailand has not used the name "Siam" for almost seventy years, it still leads a vigorous existence in Penang as the default term, where other varieties of Hokkien have adopted *Thài* 泰.
8   *Bâng-gà-lí* 芒加里 derives erroneously from *Bengali*.

meaning is still retained in some expressions,[9] it has come to refer to the deities themselves, rather than their representations in stone and clay. Without further qualification, the word mainly refers to deities worshipped by the Chinese, but with the addition of the ethnic prefix *Kè-lêng-á* 吉寧仔, it may also refer to Hindu deities. The gods of monotheistic religions are seldom referred to in this fashion, presumably because they seldom are represented with statues. Sikhism is exceptional, and Sikhs are said to worship *Bâng-gà-lí ang-kong* 芒加里尪公 'Bengali gods'. Larger Taoist temples are named *ang-kong-am* 尪公庵;[10] these are administered by groups known as *ang-kong-hōe* 尪公會 'god clubs'. Small temples or residential houses that double up as places of worship are known as *ang-kong-keng* 尪公間. The birthdays of Chinese deities according to the Lunar calendar are known as *ang-kong-se·ⁿ* 尪公生 'god's birthday' and special sacrifices are made to them on such days. Talismans or good luck charms thought to be blessed by a god are known as *ang-kong-hû* 尪公符 'divine talismans' or sometimes just as *hû*. Lucky numbers received in a dream or received from a shaman are known as *ang-kong-jī* 尪公字 'divine characters' and used for gambling and fortune-telling purposes.

In other varieties of Hokkien, *châng* 欉 is a measure word or numerical coefficient for trees and plants. In other varieties, it is termed *chhiū* 樹 or *chhiū-bók* 樹木. The Penang usage for 'tree' probably derives from an older word *chhiū-châng* 樹欉, meaning a bush or shrub, which Penangites still sometimes use as the word for tree. In compounding, Penang Hokkien takes the final syllable of this words and uses it as a general term for all types of bushes, shrubs, plants and trees, whatever their size. Thus, a coconut palm is a *iâ-châng* 椰欉, bamboo is *tek-châng* 竹欉, a banana tree is a *keng-chio-châng* 弓蕉欉, a rambutan tree is *âng-mô·-tan-châng* 紅毛丹欉, a coffee plant is a *ko-pì-châng* 嗄呸欉, and a rubber plant is *chhiū-leng-châng* 樹朥欉.

Hì 戲 originally referred to traditional theatre performances and the stage on which these are performed, and in other varieties it still retains this meaning. In Penang, however, this is the ordinary word for a film or TV programme. For the traditional type of outdoor theatre performances, the Malay loan *way-ang* is also used. Therefore, *khoàⁿ-hì* 看戲 refers to watching a movie or TV programme, rather an attending a theatre performance as it would in other varieties. An *Âng-mô·-hì* 紅毛戲 'red-hair film' is a Hollywood movie or a

---

9   In Penang, this meaning is retained with the diminutive suffix *á* 仔 and can be found in compound terms such as *ang-kong-á* 尪公仔 'a doll, a figurine, or a picture of a character or a face' and *ang-kong-á-piáⁿ* 尪公仔餅 'small cakes shaped like an animal or with a picture of an animal on them made and eaten during the Mid-Autumn festival', also known as mascot mooncakes.

10 A less commonly used alternative is *ang-kong-biō* 尪公廟.

movie from a Western country, whereas a *kúi-hì* 鬼戲 'ghost film' is a horror film. To make or film a movie is to *hip-hì* 翕戲 'to take a film', modelled on *hip siàu-éng* 翕相影 'to take a photo', and to download one from the internet – a more recent term, indicating that Penang Hokkien is still capable of this type of invention – is to *lóh-hì* 落戲 'to make a film come down'. *Hì-tâi* 戲臺 generally refers to a cinema in Penang, whereas in other varieties it usually means a theatre or stage.

In addition to its original meaning of 'a letter, word, or written character', *jī* 字 has the additional meaning of 'a certificate or official document' that is absent from modern Taiwanese or Amoy usage. Thus, a birth certificate is a *chhut-se·ⁿ-jī* 出生字,[11] a death certificate is *sí-lâng-jī* 死儂字 'dead person certificate', and a *koa-sa-jī* 啩唦字 is a will or document of power of attorney, the first part of which is borrowed from Malay *kuasa* 'power'.

Chiáh 食, aside from the meaning 'to eat' known in other varieties of Hokkien, has an additional function in Penang as a noun indicating different types of ethnic cuisine or dishes, and replaces a number of terms in other varieties of Hokkien. *Âng-mô·-chiáh* 紅毛食 is a general term for Western food, *Hoan-á-chiáh* 番仔食 refers to Malay food (*nasi Melayu*), Chinese food is *Tn̂g-lâng-chiáh* 唐儂食, and *Siām-chiáh* 暹食 refers to Thai food. This way of distinguishing Chinese, Western, Malay, and Indian food is just one example of a very large class of compounds in Penang Hokkien associated with the four main ethnic groups that the speakers of the language distinguished in their daily life, as will be discussed in the next section.

4　　"Ethnic" Compounding Prefixes *Âng-mô·*, *Tn̂g-lâng*, *Hoan-ná*, and *Kè-lêng-á*

Prior to the nineteenth century, things the Chinese saw or knew to be foreign imports were prefixed variously with 胡 (hu), 洋 (yang), 番 (fan) and so on, to indicate their foreignness. Once used in this manner, they no longer counted as adjectives, but were fixed as part of the word (Masini 1993: 124–25). As a result of being spoken in a multicultural society, Penang Hokkien has created many different words for the social and material cultures they encountered through prefixing them in different ways, depending on which groups they perceived them to be associated with. Marked categories of the "foreign" are subdivided into *Âng-mô·* 紅毛 'Western', *Hoan-á* 番仔 'Malay', and *Kè-lêng-á* 吉寧仔 'Indian', with an additional prefix *Tn̂g-lâng* 唐儂 for Chinese things that

---

11　　An alternative form is *chhut-se·ⁿ-chóa* 出生紙 'birth paper'.

were the once the norm, but in a multicultural context require a prefix to mark them out as specifically Chinese.[12] The preceding four terms are known in the geographically close varieties spoken in Singapore and Medan, but Penang Hokkien has been especially creative in using them for derived compounds.

The term *Âng-mô·* 紅毛 'red-hair' is both the general term for Westerners or Europeans, but also has a narrower meaning that refers specifically to English-speaking peoples. The word was once in widespread usage throughout East and Southeast Asia, where it generally meant either Dutch or English people depending on the colonial context (see Ding, this volume). The term is still well-known in Singapore, even in the local variety of English, and is also used in Thai and Cambodian Teochew, and its cognates are used in Borneo Hakka. *Âng-mô·* has a wide range of uses in Penang for compounding unparalleled in other varieties.[13] *Âng-mô·* was already very widely used for compounding by 1900. Lo Man Yuk records seven names that use it, including *Âng-mô·-kong-koan* 紅毛公館 'town hall', the *Âng-mô·-tōa-lé-pài-tn̂g* 紅毛大禮拜堂 'St. George's church', and *Âng-mô·-óh* 紅毛學 'Penang Free School'.

Compounds relating to people of English or Western ethnic background include *Âng-mô·-kiáⁿ* 紅毛囝 'red-hair children' for Western children, *Âng-mô·-cha-bó·* 紅毛查某 'a Western woman', *Âng-mô·-pô* 紅毛婆 'red-hair wife' as a slightly derogatory term for a Western woman, and *Âng-mô·-kâu* 紅毛猴 'red-haired monkey' as a not particularly polite way of referring to a Western man. To be or do something *âng-mô·-khóan* 紅毛款 'red-hair-style' is a Western way or style of doing something. As the language typically spoken by the type of Westerners with whom the Penang Chinese had the most frequent intercourse, *Âng-mô·* has the extended meaning of the English language as spoken or written (in Medan Hokkien it often referred to Dutch). *Âng-mô·-sái* 紅毛屎 'red-hair shit' derives from the longer expression *chiáh-âng-mô·-sái* 食紅毛屎 'to eat the red-hair shit', the literal meaning of which is to receive an English education. The short form is used as a noun to refer to someone educated in an English-medium school who knows no written Chinese, and is frequently employed jocularly or in a self-deprecatory manner.

---

12 For example *saⁿ* 衫 is the general term for clothing, and *Tn̂g-lâng-saⁿ* 唐儂衫 now refers to traditional Chinese clothing such as cheongsams. At some point in the past, when Chinese styles of dress were the norm, Western clothing would presumably have been *Âng-mô·-saⁿ* 紅毛衫.

13 A significant number of these terms existed in mid-century Malayan Cantonese as well (see Bruce 1954: 122), but these are seldom used nowadays. A few of these words still exist other varieties of Hokkien: *âng-mô·-thó·* 紅毛土 (Taiwan) or *âng-mô·-thó·* 紅毛灰 (China) for 'cement' and *âng-mô·-tan* 紅毛丹 for 'rambutan' (see also Ding, this volume).

Another use of *âng-mô·* 紅毛 is as a prefix is used to mark things or concepts associated with Western or colonial culture, often in contrast to their perceived equivalents in Chinese culture (Table 5.3). Sometimes these are used only when a distinction needs to be made, for example *âng-mô·-chhit-goéh* 紅毛七月 'red-hair seventh month' is July only when it needs to be disambiguated from the seventh month of the Chinese lunar calendar. *Âng-mô·-tang-cheh* 紅毛冬節 'red-hair's Winter Solstice', a common word for Christmas, is formed on analogy with the Chinese festival of *Tang-cheh* 冬節 'Winter Solstice', because the two festivals are only four days apart.[14] However, the term for All Souls' Day, *âng-mô·-chheng-bêng* 紅毛清明, is named after the tomb Sweeping Day or the *Chheng-bêng* 清明 festival, not because of the similar date, but because both are festivals at which the dead are remembered.

TABLE 5.3 Compounds with *Âng-mô·* 'Westerner'

English	Penang Hokkien	Meaning of Penang term in other varieties	Amoy/Taiwan equivalent
All Souls' Day	*âng-mô·-chheng-bêng* 紅毛清明	red-hair tomb-sweeping day	*tui-su í-bông-jít* 追思已亡日
Arabic numerals (as opposed to those in Chinese characters)	*âng-mô·-hō-bé* 紅毛號碼	red-hair numbers	*A-la-pek sò·-jī* 阿拉伯數字
botanical gardens	*âng-mô·-hoa-huîⁿ* 紅毛花園	red-hair flower garden	*sít-bút-hñg* 植物園
British army; British soldiers	*âng-mô·-peng* 紅毛兵	red-hair soldiers	*Eng-kun* 英軍
British colonial period	*âng-mô·-chhiú* 紅毛手	red-hair hand	*Eng-kok sit-bîn--ê sî-tài* 英國殖民个時代
bungalow	*âng-mô·-lâu* 紅毛樓	red-hair building	(no exact translation)
Christmas	*âng-mô·-tang-cheh* 紅毛冬節	red-hair winter solstice	*Sèng-tàn* 聖誕, *Sèng-tàn-chiat* 聖誕節

14 Alternatives are *Ia-So·-se·ⁿ* 耶穌生 'Jesus' birthday' and *Hōng-kàu-se·ⁿ* 奉教生 'Catholicism birthday'.

TABLE 5.3   Compounds with *Âng-mô·* 'Westerner' (*cont.*)

English	Penang Hokkien	Meaning of Penang term in other varieties	Amoy/Taiwan equivalent
English first name (either official or unofficial, used in addition to or in place of one's Chinese name)	*âng-mô·-miâ* 紅毛名	red-hair name	*Eng-bûn-miâ* 英文名
English language	*âng-mô·(-ōa)* 紅毛(話)	red-hair language	*Eng-gú ~ Eng-gí* 英語
Roman alphabet, Roman letters, written English	*âng-mô·-jī* 紅毛字	red-hair characters	*Lô-má-jī* 羅馬字, *Peng-im* 拼音, *Eng-bûn* 英文
solar calendar (New Year beginning on 1 January)	*âng-mô·-chiaⁿ* 紅毛正	red-hair first month	*goân-tàn* 元旦 (Amoy), *sin-nî* 新年 (Taiwan)
solar calendar year	*âng-mô·-nî* 紅毛年 (also used in Singapore)	red-hair year	*iông-lék-nî* 陽曆年
sweets	*âng-mô·-thn̂g* 紅毛糖	red-hair sugar	*thn̂g-á* 糖仔, *thn̂g-kó* 糖果
tomato	*âng-mô·-kiô* 紅毛茄	red-hair brinjal	*tho-ma-toh* (Taiwan), *chhàu-khī-á* 臭柿仔 (Amoy)
Western age (when it needs to be disambiguated for official purposes from the Chinese age that adds an extra year)	*âng-mô·-hòe* 紅毛歲	red-hair years	
Westerner, white foreigner; the English language, English as a school subject	*âng-mô* 紅毛 (also used in Singapore)	red hair	*gōa-kok-lâng* 外國人, *a-tok-á* 阿啄仔 (Taiwan), *hoan-á* 番仔 (Amoy), *Eng-gú* 英語, *Eng-bûn* 英文

The last four terms are connected specifically to things associated with the British colonial period and require some historical context. *Âng-mô·-lâu* 紅毛樓 'red-hair building' refers to the kind of luxurious bungalow that Westerners owned and lived in in colonial times, and does not correspond neatly to any term in Taiwanese or Amoy Hokkien. The word is also used in a jocular sense as a term for a bump on the head received from a blow. *Âng-mô·-hoa-huîⁿ* 紅毛花園 'the red-hairs' garden' originally referred specifically to the Botanical Gardens at Air Itam, but through association this term ended up being used by Penangites overseas to translate the English names for any botanical gardens. *Âng-mô·* remains a productive prefix in Penang Hokkien; a brand of craft beer recently released in Penang was humorously named "Ang Moh Leng Te" 紅毛冷茶 'the red-hairs' cold tea'.

Tn̂g-lâng 唐儂 'Tang person' is the second most common of the ethnic prefixes, indicating Chinese ethnicity, and is generally used for those who do not hold Chinese nationality. The term predates modern Chinese concepts of nationalism, whilst still recognizing commonalities of culture and language with other Chinese groups such as Cantonese and Hakkas.[15] The related terms *T'ɐ̄ng-yān*, *Tong-nyin*, and *Tn̂g-nâng* exist in spoken Cantonese, Hakka, and Teochew respectively, but *T'ɔ̄ng-yān* is generally no longer used amongst younger Cantonese speakers in Malaysia, for whom *Wā-yān*, modelled on Mandarin *Huaren* 華人, is the preferred term. A feature of Penang-style Hokkien is the use of the disyllabic *Tn̂g-lâng* 唐儂 in the creation of new compound words, where earlier usage preferred the monosyllable *Tn̂g* 唐. Barclay's *Dictionary* (1873) and Francken and de Grijs' *Dictionary* (1889) both list *Tn̂g-lâng* with the meaning 'a Chinese', but list only the compounds *Tn̂g-ōa* 唐話 for the Chinese language and *Tn̂g-jī* 唐字 for Chinese characters,[16] and Barclay's supplement of 1923 does not contain any new compounds with either *Tn̂g* or *Tn̂g-lâng*. In contrast, Penang Hokkien has many compounds prefixed with *Tn̂g-lâng* to mark connection with Chinese material and social culture (Table 5.4). In this way, Penang Hokkien speakers differentiate them from the similar features in the foreign cultures to which Chinese in the Straits Settlements were exposed.

---

15   The name *Tn̂g-lâng* first appears in a Southeast Asian context in Ma Huan's 馬歡 1433 description of Java his book *The Overall Survey of the Ocean's Shores* (瀛涯勝覽 *Yingyai shenglan*) as a collective term for people from Guangdong, Zhangzhou, Quanzhou, and other places who have come live in Java, some of whom had converted to Islam.

16   Medhurst (1832: 661) has both the literary (*Tông-jîn*) and colloquial (*Tn̂g-lâng*) reading.

TABLE 5.4 Compounds with *Tn̂g-lâng* 'Chinese'

English	Penang Hokkien	Meaning of Penang term in other varieties	Amoy/Taiwan equivalent
age according to Chinese reckoning (usually one's Western age plus one)	*Tn̂g-lâng-hòe* 唐儂歲	Tang person age	*hòe* 歲
birthday according to the lunar calendar	*Tn̂g-lâng-se·ⁿ-jít* 唐儂生日	Tang person birthday	*long-lék se·ⁿ-jít* 農曆生日
Chinese characters	*Tn̂g-lâng-jī* 唐儂字	Tang person characters	*Hàn-jī* 漢字
Chinese doctor, doctor of TCM	*Tn̂g-lâng-sin-se·ⁿ* 唐儂先生	Tang person master	*tiong-i* 中醫
Chinese language; Mandarin	*Tn̂g-lâng-ōa* 唐儂話	Tang person speech	*Tiong-bûn* 中文, *Tiong-kok-ōe* 中國話, *Hôa-gú* 華語, *Kok-gí* 國語, *Phó·-thong-ōe* 普通話
Chinese lunar calendar	*Tn̂g-lâng-láh-jít-tô·* 唐儂曆日圖	Tang person calendar	*kū-lék* 舊曆, *lông-lék* 農曆
Chinese-medium school	*Tn̂g-lâng-óh(-tn̂g)* 唐儂學(堂)	Tang person school	*Tiong-bûn hák-hāu* 中文學校
Chinese name	*Tn̂g-lâng-miâ* 唐儂名	Tang person name	*Tiong-bûn-miâ* 中文名
Chinese New Year	*Tn̂g-lâng-chiaⁿ* 唐儂正	Tang person first month	*sin-chiaⁿ* 新正, *kòe-nî* 過年, *chhun-chiat* 春節 (Amoy)
Chinese person, Overseas Chinese; Chinese (language), Chinese (school subject)	*Tn̂g-lâng* 唐儂	Tang person	*Hàn-chók* 漢族 (Amoy), *Hôa-jîn* 華人, *Tiong-bûn* 中文, *Tiong-kok-ōe* 中國話, *Hôa-gú* 華語, *Phó·-thong-ōe* 普通話
Chinese tea	*Tn̂g-lâng-tê·* 唐儂茶	Tang person tea	*tê* 茶

TABLE 5.4   Compounds with *Tn̂g-lâng* 'Chinese' (*cont.*)

English	Penang Hokkien	Meaning of Penang term in other varieties	Amoy/Taiwan equivalent
third month of the lunar calendar	*Tn̂g-lâng san-goéh* 唐儂三月	Tang person March	*long-lék sa-goéh* 農曆三月
to be conservative in one's attitudes; to have old-fashioned Chinese ideas	*Tn̂g-lâng-sim* 唐儂心	Tang person heart	*pó-siú* 保守
to study in a Chinese-medium school	*thák Tn̂g-lâng-chhe·h* 讀唐儂冊	read Tang person books	*thák tiong-bûn hák-hāu* 讀中文學校 (Taiwan), *thák tiong-bûn óh-tn̂g* 讀中文學堂
Traditional Chinese Medicine	*Tn̂g-lâng-ióh* 唐儂藥	Tang person medicine	*tiong-ióh* 中藥, *Hàn-ióh* 漢藥

In the past, *Tn̂g-Soan* 唐山 (lit. 'Tang Mountain') was the common vernacular word for 'China' in almost all ethnic Chinese communities outside China, who had migrated from the southern coast of China. The term has the connotations of the "old country" and in Penang, those who had been born in China and spent a portion of their life there were known as *Tn̂g-Soan-lâng* 唐山儂. An additional term, *Tn̂g-Soan a-pe·h* 唐山阿伯 (lit. 'Tang Mountain uncle') refers to the older generation of men who were born and grew up in China before coming to Malaya, and to act *Tn̂g-Soan--lâi* 唐山來 (lit. 'coming from Tang Mountain') is used for perceived unfavourable characteristics or conservative behaviour associated with these people. It is only recently that *Tn̂g-Soan* 唐山 for 'China' has become archaic and only older people use it, whereas people under the age of sixty would use *Tiong-Kok* 中國.[17] Nevertheless, the distinction between a *Tn̂g-lâng* and a *Tiong-Kok-lâng* 中國儂 is still maintained, in

---

17   This may be because of its associations with the expression *túin Tn̂g-Soan* 轉唐山 'to return to China' being used as a euphemism for death.

the way that *Huaren* 華人 'an ethnic Chinese' and *Zhongguoren* 中國人 'a PRC national' are marked off in Malaysian Mandarin.

The term *Hoan-á* 番仔 is used in all varieties of Hokkien, but has differing meanings depending on where Hokkien is spoken. Its most basic meaning lies somewhere between 'savage' and 'foreigner', but in different varieties the meaning has narrowed and became attached to specific groups of people, usually the indigenous peoples in the lands to which Hokkiens have migrated. In Taiwanese it refers to Formosan aboriginals,[18] in the Philippines generally to Tagalogs, in Thailand to native Thais in contrast to Thai Chinese, and in Amoy – where there are no indigenous people other than the locals – it refers to Westerners, substituting the last syllable of the older term for Westerners *âng-mô·-hoan* 紅毛番 for the whole word, where other varieties have dropped it. *Hoan-á* was formerly used in both Amoy and Taiwan with a wider meaning, marking things that had come from overseas, such as *hoan-á-hóe* 番仔火 'matches'. In Penang, the word has narrowed its meaning and refers specifically to Malays. It is used in a similar way to the above two terms: to mark out things associated with Malay culture. The offensive "savage" connotation of the word has generally been lost. In contrast to Amoy Hokkien, where the word is falling out of use, *Hoan-á*, like *Âng-mô·*, has taken on a life of its own. Thus a *Hoan-á-kiáⁿ* 番仔囝 'Malay son' is a young Malay man, a *Hoan-á-khu* 番仔區 is an area where Malays form the majority of the population, and *Hoan-á-chiáh* 番仔食 refers to Malay cuisine (*nasi Melayu*). The word *Hoan-á* also refers to the Malay language as spoken or as a subject taught at school. To *kóng Hoan-ná* 講番仔 is to speak Malay, and *Hoan-á-chhe·h* 伊番仔冊 (lit. 'Malay books') refers to Malay-medium schooling. As the Roman (*Rumi*) script has long replaced Arabic-based Jawi as the most common mode of writing Malay, the term *Hoan-á-jī* 番仔字 (lit. 'Malay words') refers to written Malay in roman letters, whereas the older Arabic-based Jawi script – still a fairly common sight in Malaysia – is termed *tāu-gê·-jī* 豆芽字 'bean sprout letters' because of its resemblance to bean sprouts. The religion of the Malays, Islam, although generally termed *Hôe-kàu* 回教, is also referred to with *Hoan-á* in a few compounds (Table 5.5).

---

18 This usage is now considered offensive, and the neutral *goân-chú-bîn* 原住民 is now the preferred term.

TABLE 5.5   Compounds with *Hoan-á* 'Malay'

English	Penang Hokkien	Meaning of Penang term in other varieties	Amoy/Taiwan equivalents
Eid, Hari Raya	*Hoan-á-chia^n* 番仔正	[not understood]	*Khai-chai-cheh* 開齋節
Malay person; Malay language	*Hoan-á* 番仔	Westerner (Amoy), Indigenous Formosan (Taiwan)	*Má-lâi-lâng* 馬來儂
Mosque	*Hoan-á-am* 番仔庵	[not understood]	*chheng-chin-sī* 清真寺
to convert to Islam	*jíp-Hoan* 入番[a]	[not used, could be taken to mean to marry into an indigenous family in Taiwan]	*kui-hiòng hôe-kàu* 歸向回教 (Taiwan)
to speak Malay	*kóng hoan-ná* 講番仔	to speak English (Amoy),[b] to speak an indigenous Formosan language (Taiwan)	*kóng Má-lâi-ōe* 講馬來話

a   A calque from Malay *masuk Melayu* in the same meaning.
b   Taiwanese and Amoy would demand that the *ōe* 話 'language' suffix is used. In Penang Hokkien, this is pronounced *ōa* and is optional.

The last of the four most common ethnic prefixes is not a native Hokkien word, but is noted here as it functions in the same way as the above three. This is *Kè-lêng-á* 吉寧仔, usually translated carelessly into English as 'Indian', which refers specifically to Tamil-speaking Hindu Malaysians of Indian descent, and derives from the Malay *Keling*. This word ultimately derives from Sanskrit *Kāliṅga*, the name of an old kingdom on the Coromandel Coast mentioned in the second chapter of the *Sejarah Melayu* 'Malay Annals'. Sensitivities around its use – due to the erroneous belief that the term originates from the onomatopoeic "clink-clink" of either Indian foot-bangles or the chains of indentured labourers – have resulted in some speakers using *Ìn-tō·* 印度 to replace it in recent years. However, this word makes no distinction between Indian nationals and Malaysian Indians, and also fails to take into account the linguistic and cultural differences amongst different groups of Indians,

as *Kè-lêng-á* in most contexts refers to Tamil speakers specifically. Therefore *Kè-lêng-á-ōa* 吉寧仔話 is the Tamil language, and it is written in *Kè-lêng-á-jī* 吉寧仔字 (lit. 'Kling characters'), the Tamil script. Some exceptions to this are the terms *Kè-lêng-á-hì* 吉寧仔戲, which refers to any Indian films, be they from Bollywood or Kollywood, *Kè-lêng-á-hióh* 吉寧仔箬 'Indian leaves' for curry leaves, and terms relating to the Hindu religion. Hindu gods and notable Hindu temples in Penang are known as *Kè-lêng-á ang-kong* 吉寧仔尫公, although *Kè-lêng-á-am* 吉寧仔庵 is also used in the latter meaning. To pray to these deities or *pài Kè-lêng-á ang-kong* 拜吉寧仔尫公 also has the extended meaning of being a Hindu. *Kè-lêng-á-chia^n* 吉寧仔正 'Indian New Year' is the festival of Deepavali, although – like *Hoan-ná-chia^n* (Hari Raya) – this festival is unconnected to the advent of a new year; *chia^n* 正 has undergone a shift in meaning to indicate the most important festival in the religious calendar rather than a new year (Li 2007: 61–63). Another term related to food, *Kè-lêng-á-pūi^n* 吉寧仔飯, refers to Indian-style curries served with steamed rice, but is also used in a similar way to English "porridge": as a metaphor for prison food. To say someone is *chiáh-Kè-lêng-á-pūi^n* 食吉寧仔飯 'eating Indian curries' is equivalent to saying they are "doing time".

5 Local Concepts and New Things

A number of compound terms exist for things that were rare or unknown in late nineteenth-century Fujian, and therefore usually have no Amoy or Taiwanese equivalent. These relate to the history of Malaysia, the colonial government, local industries, and names of subvarieties of local fruits (Table 5.6).

TABLE 5.6   Words for local politics, officialdom, local industries, and material culture

English	Penang Hokkien	Meaning
area under government control during the Malayan emergency	*pé·h-khu* 白區	white district
communist-controlled area during the Malayan emergency	*o·-khu* 烏區	black district
dredge for a tin mine	*thih-chûn* 鐵船	ironclad
formic acid (used to coagulate rubber)	*chhiū-leng-chhò·* 樹�롄醋	rubber vinegar

TABLE 5.6   Words for local politics, officialdom, local industries, and material culture (*cont.*)

English	Penang Hokkien	Meaning
governor of Penang	*jī-ông* 二王	second king; viceroy
governor's residency	*jī-ông-chhù* 二王厝	second king's house
I.C. office (where one goes to pick up one's national identity card)	*teng-kì-koan* 登記關	registration office
kebaya (the embroidered blouse traditionally worn by the Nyonyas)	*poàⁿ-tńg-té* 半長短	half-long-short (a description of the garment's uneven length)
Malaysian Federal Reserve Unit	*âng-thâu-peng* 紅頭兵	the red head soldiers (the name derives from their red berets)
mangle (used for turning out sheets of rubber)	*chhiū-lêng-ká* 樹朥絞	rubber mangle
opencast mine; tin mine	*khut-lông* 窟廊	pit corridor
orang minyak (a type of Malay ghost covered with oil)	*o·-iû-lâng* 烏油儂, *o·-iû-kúi* 烏油鬼	oil man (calque from Malay); oil ghost
pisang emas	*kim-chio* 金蕉	gold banana (calque from Malay)
pisang keling (a type of short, sweet banana)	*Kè-lêng-á-chio* 吉寧仔箸	Indian banana
police detective	*âm-pâi* 暗牌 (also known in Singapore)	hidden badge (on account of a detective wearing no uniform)
police sergeant	*saⁿ-liáp-che ⁿ* 三粒星, *saⁿ-oáh* 三劃	three stars; three stripes
rubber plantation	*chhiū-leng-pa* 樹朥芭	rubber forest (*pa* is a probable loan from Thai *paa* ป่า)
smaller, wild variety of durian	*soaⁿ-liû-liân* 山榴槤	mountain durian
soursop (*Annona muricata*)	*âng-mô·-liû-liân* 紅毛榴槤	red-hair durian (presumably a calque from Malay *durian belanda*, also known in Singapore)
tin refinery	*siah-bí-lông* 錫米廊	tin ore corridor

TABLE 5.6   Words for local politics, officialdom, local industries, and material culture (*cont.*)

English	Penang Hokkien	Meaning
to tap rubber	*koah-chhiū-leng* 割樹朥, *phah-chhiū-leng* 撲樹朥	to cut rubber; to hit rubber
traffic police	*péh-kha-té* 白骹底 (also known in Singapore)	white leg-bottoms (a reference to the white spats that are part of some police uniforms)
type of durian with a particularly green rind	*chhe·ⁿ-phôe-kiáⁿ* 青皮囝	green-skinned son
type of durian with an orangey flesh	*âng-bah-kiáⁿ* 紅肉囝	red-fleshed son
type of moist banana	*tâ-lî-chio* 銅鐳蕉	copper coin banana

Penang Hokkien speakers required new words for inventions and institutions that became common throughout the industrialized and urbanized world of the late nineteenth and twentieth centuries. They were open to borrowing to fill these lexical gaps, and many terms such as 'taxi', 'radio', 'bus', 'cheque', 'gas tank', 'license', 'lorry', 'tyre', and 'commission' entered the language, modified to some extent to fit to Hokkien phonology. Sino-Japanese words for some of these new things – such as *tiān-ōa* 電話 'telephone' – did make their way into Penang Hokkien, as did some later coinages for things that became common in the last decades of the twentieth century, such as computers and air conditioning. These were adopted into Penang Hokkien during the 1970s or 1980s. At a time when the ability to read and write Mandarin Chinese was already on the rise in Malaysian Chinese communities, there was also increased consumption of entertainment from Taiwan and Hong Kong. As the Mandarin used outside China generally followed ROC national standard of *Guoyu* – in preference to *Putonghua* promulgated as the standard within the PRC – borrowings dating from this time reflect this trend. Penang Hokkien still uses ROC *léng-khì* 冷氣 'air conditioning' and *tiān-náu* 電腦 'computer' in place of PRC coinages such as *khong-tiâu* 空調 and *kè-sǹg-ki* 計算機, which are the preferred terms in Amoy. Aside from these types of loanwords, Penang Hokkien speakers have derived a wide range of neologisms that are incomprehensible to speakers of other varieties, or misunderstood because they have different meanings (Table 5.7).

TABLE 5.7 Words for new inventions

English	Penang Hokkien	Meaning of Penang term in other varieties	Amoy/Taiwan equivalents
ambulance	*âng-síp-jī-chhia* 紅十字車, *âng-cháp-jī-chhia* 紅十字車,[a] *ló-kun-chhia* 老君車	red cross vehicle, doctor vehicle	*kiù-hō·-chhia* 救護車
bank account	*hō·-kháu* 戶口	household (Taiwan), household registration (Amoy)	*kháu-chō* 口座 (Taiwan), *siàu-hō* 數號 (Amoy)
crash helmet	*thih-bō* 鐵帽	iron hat	*an-choân-bō* 安全帽
electrical socket	*tiān-thâu* 電頭	electricity head	*chhah-chō* 插座
holiday house	*chiáh-hong-lâu* 食風樓[b]	eat wind building	*piát-sū* 別墅
hospital	*ló-kun-chhù* 老君厝, *ló-kun-lâu* 老君樓,[c] *pē·ⁿ-chhù* 病厝	doctor house, doctor building, sick house	*pēⁿ-īⁿ* 病院, *i-īⁿ* 醫院
life ring	*pó-ke·-kho·* 保家箍	insurance ring	*kiù-seng-khoân* 救生圈
lifejacket	*pó-ke·-saⁿ* 保家衫	insurance jacket	*kiù-seng-i* 救生衣
maternity hospital; maternity ward	*se·ⁿ-kiáⁿ-tn̂g* 生囝堂	hall for giving birth	*hū-sán-kho i-īⁿ* 婦產科醫院

a Red crosses have not been used on ambulances since the Malaysian red cross became Malaysian Red Crescent in September 1975, but the word is still widely known and used.

b *chiáh-hong* 食風 'to eat the wind' is a calque from Malay *makan angin*, with the extended meaning of taking the air or taking a holiday.

c Both *pē·ⁿ-chhù* 病厝 and *ló-kun-chhù* 老君厝 appear in Lo Man Yuk as Penang usage, but *pē·ⁿ-chhù* 病厝, possibly a calque from older Malay *rumah sakit*, appears to have largely fallen out of use.

TABLE 5.7   Words for new inventions (*cont.*)

English	Penang Hokkien	Meaning of Penang term in other varieties	Amoy/Taiwan equivalents
mental hospital; lunatic asylum	*siáu-lâng-keng* 痟儂間	crazy person house	*sin-keng-pēⁿ-īⁿ* 神經病院
mobile phone	*chhiú-tiān* 手電, *chhiú--ê tiān-ōa* 手个電話	torch, hand telephone (calque from Malaysian English *handphone*)	*chhiú-ki* 手機
plastic bag	*goân-chú-lông* 原子囊, *goân-chú-tē* 原子袋	particle bag (*plastic-lông* is also commonly used)	*sok-ka-tē* 塑膠袋 (Taiwan), *sok-liāu-tē* 塑料袋 (Amoy)
refrigerator	*sng-tû* 霜櫥	ice box	*peng-siuⁿ* 冰箱
traffic roundabout	*îⁿ-kho·-khoân* 圓箍圈	circle-ring-circle	*îⁿ-khoân* 圓圈 (Taiwan), *khoân-tó* 環島 (Amoy)
transformer; diesel generator; car battery; electricity meter	*tiān-siauⁿ* 電箱	electricity box	*piàn-ap-khì* 變壓器, *chhâ-iû hoat-tiān-ki* 柴油發電, *tiān-pió* 電表

Several terms are derived through semantic extension, whereby a pre-existing word is used in a novel way unknown elsewhere. Some of these new usages derive from perceived resemblances, and others to well-known advertising symbols associated (or formerly associated) with the extended term. These words often retain their original meanings in addition to the extended idiomatic usage (Table 5.8).

Aside from these everyday expressions, Penang Hokkien contains a large number of slang and argot terms derived through idiomatic extension related to sexual activity, death, or other taboo subjects, such criminal as triad-related activities. These words deserve their own separate study.

TABLE 5.8   Extended meanings

Penang Hokkien word	Original meaning	Extended meaning
âng-hê· 紅蝦	red prawn	a cultivar of durian with an orangey flesh
*bong-kha-chhui*ⁿ 摸尻川	to touch a bum	a glove puppet (so named because the hand goes up the rear of the puppet)
*chhiāu*ⁿ*-gê·* 象牙	elephant tusk; ivory	a type of large, greenish banana
*chiàu-kià*ⁿ 照鏡	to reflect in the mirror	to undergo an X-ray
hê·-bóe 蝦尾	prawn tail	an outboard motor, a propeller (named on account of its appearance; in Taiwan *thui-chìn-khì* 推進器)
*hóe-chì*ⁿ 火箭	rocket	the Democratic Action Party (DAP; from the rocket used as the party's symbol)
lāu-hó·-thâu 老虎頭	tiger's head	banknote; paper currency (from a series of Malaysian banknotes that had a tiger's head as the watermark, removed from 1982 onwards)
liām-keng 念經	to chant a Buddhist sutra	to attend Friday prayers as a Muslim; to read the Quran; to be a practising or devout Muslim
o·-káu 烏狗	black dog	stout, Guinness (so named because the brand of stout sold in Malaysia for many years had a bulldog's head on the bottle)
pùn-chhìn 磅秤	pair of scales	Barisan National (so called because the party uses a pair of scales as its symbol)
*soa*ⁿ*-téng* 山頂	on the mountain	out of town; in the suburbs; suburban
tu-lông 豬櫳	pigpen	a child's playpen

6   Concluding Remarks

The terms highlighted in this chapter reflect a common tendency of Sinophone Southeast Asia towards creating their own terms rather than simply borrowing from Mandarin. These locally invented elements, often from purely Hokkien components, demonstrate the importance of "thinking beyond *rojak*". They

also provide a case study in which Chineseness is problematized, as many Penang Hokkiens can express their identity in a variety of ways. Nevertheless, over the past twenty or so years, conditions amenable to the development of distinctive neologisms have greatly altered for Penang Chinese. Although there is considerable community interest in Penang Hokkien – resulting in the publications of dictionaries, collections of songs and poetry, a revival campaign, as well as the production of a weekly podcast and a feature film in the language – the ability to speak, understand, read, and write Mandarin has increased significantly since the 1980s to the point that many Penang Chinese under the age of thirty are no longer proficient in speaking Penang Hokkien. Through increased familiarity with written Chinese, Mandarin, and the speech habits of Taiwanese Hokkien speakers, many younger speakers have begun to modify their vocabulary in accordance with their knowledge of these other languages, because they perceive these speech forms to be "purer" and more "correct" than the *rojak* variety spoken by their parents and grandparents. These attitudes, coupled with ethnic tensions within Malaysia, have resulted in a sharp decrease in the use of Malay loanwords and a new layer of superstrate Mandarin vocabulary entering Penang Hokkien. They do so either as direct loans in which the Mandarin pronunciation remains intact (for instance *Huayu* 華語 for 'Mandarin'), or as graphic loans or loans by analogy in which the Mandarin compound word generates its real or imagined cognate in Hokkien according to its characters (e.g. the same characters pronounced as *Hôa-gú*).

Hokkien programming on Hua Hee Dai presents a language that is still recognisable as Penang Hokkien due to its pronunciation, but reflects all of the above trends of following Mandarin usage as a standard and is fairly restricted in its use of Malay and English loanwords. The consequences of these trends for distinctive Penang vocabulary is twofold: some terms may be unknown to younger speakers and end up being replaced by new formations modelled on Mandarin, such as *Tiong-i* 中醫 'Chinese medicine' for *Tn̂g-lâng-ióh* 唐儂藥, *chò-ài* 做愛 'to make love' in place of *kiân-pâng* 行房, and *Hàn-jī* 漢字 'Chinese characters' in place of *Tng-lâng-jī* 唐儂字. Other fairly well-known terms are discarded because they are considered old-fashioned, rustic, or are seen to reflect ignorance, provincialism, or old attitudes, resulting in the replacement of *Tn̂g-soan* 唐山 'Tang Mountain' by *Tiong-kok* 中國 'China' and *kè-lêng-á ang-kong* 吉寧仔尫公 'Hindu deities' by *Ìn-Tō·-kàu* 印度教 'Hinduism'. Vocabulary that is unlikely to change in this way is that related directly to Malaysian life and material culture.

Certain historical factors have led to the formation of distinctive native vocabulary in Penang Hokkien, including low Chinese literacy, a multilingual

and multicultural environment, the lack of entertainment media, and the relative isolation of Penang Chinese from other large communities of Hokkien speakers. Because of the high proportion of Malay and English loanwords in Penang Hokkien, neologisms derived from the native Hokkien element have not received much in the way of serious analysis or study. Currently, Penang Hokkien is undergoing a shift in its vocabulary due to intense contact and competition with Mandarin in education, entertainment, and social spheres. If Penang Hokkien survives as a spoken language in the future, and Mandarin Chinese remains the primary language of education for Penang Chinese, it is likely that many of these distinctive terms will soon be discarded and lost, so there is no better time than the present to collect and analyse them for the insights they provide into the history, culture, and mindset of the speakers of this distinctive language.

References

Barclay, Thomas, *Supplement to Dictionary of the vernacular or spoken language of Amoy* (Shanghai: The Commercial Press, 1923).

Batu Gantong, *Chrita dahulu-kala namanya Sam Kok atau tiga negri ber-prang Siok, Gwi, sama Gor di jaman Han Teow* (25 vols., Singapore: Kim Sek Chye Press, 1892).

Bruce, R., *Cantonese Lessons for Malayan Students* (Kuala Lumpur: Charles Grenier and Son, 1954).

Chan Boon Kim, *Chrita Da-hulu-kala pasal Kow Chey Thian man-gawal-kan Tong Tye Chu pergi di negri She Thian c'hu keng* (8 vols., Singapore: Fang Heng Printing Press, 1911–1913).

Chen Ronghan 陳榮翰, 'Jinxiandai Zhangzhouhua de bianhua 近現代漳州話的變化', *Zhangzhou shiye daxue xuebao* 漳州職業大學學報, 4 (1999): 80–87.

Chen Zhengtong 陳正統, *Minnanhua Zhangqiang cidian* 閩南話漳腔辭典 (Beijing: Zhonghua Shuju, 2006).

Cheng, Robert, 'Borrowing and Internal development in lexical change', *Journal of Chinese Linguistics* 15/1 (1987): 105–31.

Chuang Ching-ting, Chang Yueh-chin and Hsieh Feng-fan, 'Complete and not-so-complete tonal neutralization in Penang Hokkien', in *Proceedings of the International Conference on Phonetics of the Languages in China* (2013): 54–7.

Churchman, Catherine, 'The Eclectic Nature of Penang Hokkien Vocabulary and its Implications for Character Writing', *Journal of the Taiwanese Vernacular*, 9/1 (2017): 86–125.

de Gijzel, Luc, *Penang Hokkien Dictionary* (*English-Hokkien*) (Penang: Areca Books, 2013).

Douglas, Carstairs, *Chinese-English dictionary of the vernacular or spoken language of Amoy* (London: Trübner and Co, 1873).

Erancken, J.J.C. and de Grijs, C.F.M., *Chineesch-Hollandsch Woordenboek van het Emoi Dialekt* (Batavia: Landsdrukkerij, 1882).

Gao Ran 高然, 'Yinni Sumandala beibu de minnnan fangyan' 印尼蘇門答臘北部的閩南方言, in Li Rulong (ed.), *Dongnanya Huaren Yuyan Yanjiu* 東南亞華人語言研究 (Beijing: Beijing yuyan wenhua daxue chubanshe, 2000): 165–94.

Hing, Jia Wen, 'The Polyfunctional focus particle pun in Penang Hokkien: A Contact Perspective', in *Papers from the Chulalangkorn International Student Symposium on Southeast Asian Linguistics 2017, JSEALS Special Publication 2* (2017): 51–58.

Jones, Russell, *Loanwords in Indonesian and Malay* (Leiden: KITLV Press, 2007).

Jones, Russell, 'The Chiangchew Hokkiens, the True Pioneers in the Nanyang', *Journal of the Malaysian Branch of the Royal Asiatic Society* 82/2 (2009): 39–66.

Klöter, Henning, *Written Taiwanese* (Wiesbaden: Harrossowitz Verlag, 2005).

Kwok, Raymond, *Hokkien Baba Sayings for all Occasions* (Penang: Raymond Kwok, 2005).

Li Rulong 李如龍 and Xu Ruiyuan 徐睿淵, 'Xiamen fangyan cihui yibaiduonian lai de bianhua 廈門方言詞彙一百多年来的變化', *Xiamen daxue xuebao* (*zhexue shehui-kexueban*), 1 (2007): 84–91.

Li Yongqiu (Lee Eng Kew) 李永球, *Zi yan zi yu* 字言字語 (Penang: Borderless Trans Enterprise, 2007).

Lim, Beng S., 'Penang Peranakan Hokkien: A sociological reflection of a hybrid tongue', in Sohaimi Abdul Aziz (ed.), *Penang Peranakan Chinese and Chinese Muslims: A historical and cultural journey* (Penang: Persatuan Karyawan Pulau Pinang, 2010).

Lim, Beng S. and Teoh, Boon S., 'Malay lexicalized items in Penang Peranakan Hokkien', in Mark Alves, Pauls Sidwell and David Gil (eds.), *SEALS VIII: Papers from the 8th meeting of the Southeast Asian Linguistics Society* (Canberra: Pacific Linguistics, 2007): 148–65.

Lo Man Yuk, 'Chinese Names of Streets in Penang', *Journal of the Straits Branch of the Royal Asiatic Society, Singapore* (1900): 196–247.

Maryknoll Language Service Centre, *Maryknoll English-Taiwanese Dictionary*, 2013. http://www.taiwanesedictionary.org.

Masini, Federico, 'The Formation of Modern Chinese Lexicon and its Evolution toward a National Language: The Period from 1840 to 1898' (*Journal of Chinese Linguistics* Monograph Series 6, 1993).

Medhurst, W.H., *Dictionary of the Hok-këèn Dialect of the Chinese Language according to the reading and colloquial idioms* (Macao: East India Company Press, 1832).

Morgan, Evan, *Chinese New terms and Expressions with English Translations, introduction and Notes* (Shanghai: Kelly and Walsh Ltd., 1913).

Newbold, T.J., *Political and Statistical Account of the British Settlements in the Straits of Malacca* (2 vols., London: John Murray, 1839).

Ooi Kok Hin and Tan, Julia, 'Penang Hokkien on Life Support', *Penang Monthly* (May 2017) https://penangmonthly.com/article.aspx?pageid=7412&name=penang _hokkien_on_life_support.

Shellabear, William Girdlestone, 'Baba Malay', *Journal of the Straits Branch of the Royal Asiatic Society*, 65 (1913): 49–63.

Sim, Tze wei, 'Why are the Native Languages of the Chinese Malaysians in Decline?', *Journal of the Taiwanese Vernacular*, 4/1 (2012): 62–95.

Soon, Samantha S.M., *Tonal assignment of Malay loanwords in Penang Hokkien* (MA thesis, Nanyang Technological University, 2014).

Tan Choon Hoe, *Learn to Speak penang Hokkien Dialect* (Penang: Tan Choon Hoe, 2001).

Tan Choon Hoe, *Penang Hokkien Dialect for Penangites and Tourists* (Penang: Tan Choon Hoe, 2008).

Tan Choon Hoe, *Penang Hokkien Dialect Conversations* (Penang: Tan Choon Hoe, 2010).

Tan Siew Imm, *Penang Hokkien-English Dictionary* (Kuala Lumpur: Sunway Education group, 2016).

Wu Wenxin 吳文芯, 'Malaixiya "Bincheng Fujianhua" tezhengci yanjiu 馬來西亞"檳城福建話"特徵詞研究', *Quanzhou shifanxuyuan xuebao* 泉州師範學院學報, 32/1 (2014): 66–70.

Xu Jitun 許極燉, *Taiwanhua tonglun* (Taipei: SMC Publishing, 2000).

Ye Qianqing (Yip Xhien Ching) 葉千菁, *Malaixiya Bincheng Minnanyu yinyun yanjiu* 馬來西亞檳城閩南語音韻研究 (*A Study on the Phonology of Southern Min in Penang, Malaysia*) (MA thesis, National Taiwan University, 2014).

Zhou Changji 周長楫 (ed.), *Minnanhua dacidian* 閩南話大詞典 (Fuzhou: Fujian renmin chubanshe, 2006).

Zhou Changji 周長楫, *Minnanhua gaishuo* 閩南話概說 (Fuzhou: Fujian renmin chubanshe, 2010).

Zhou Changji 周長楫 and Zhou Qinghai 周清海, *Xinjiapo Minnanhua gaishuo* 新加坡閩南話概說 (Xiamen: Xiamen daxue chubanshe, 2000).

A Preliminary Study of *Kaoka* 高甲 Playscripts in the Philippines

Caroline Chia

1 Introduction

As part of a book on Sinitic languages in Southeast Asia, particularly those that are increasingly endangered, this chapter emphasizes "Sinophone" and "Sinitic voices" in the Philippines. Briefly defined, the Sinophone landscape involves Sinitic languages, including Hokkien, and the associated cultures and communities which historically experienced colonialism and have increasingly been marginalized in more recent times.[1] Here I focus on Sinophone speech, including the soundscapes in which speech is embedded, through theatrical performance in Hokkien. Kaoka was once a popular form of entertainment in Southeast Asia, but the Philippines is the only country in the region that still performs this theatrical art today. I furthermore aim to highlight the diversity of Sinophonic representations as seen in Kaoka playscripts. In these sources, only the phonetic elements have been preserved, whereas the logographic representations (known as Sinographs or *Hanzi* 漢字) have been omitted. As regards the "Sinitic voices" that this chapter – and this book in general – aims to highlight in the context of Southeast Asia's Chinese minorities, the academic focus has previously been on the migration and economic development of Chinese communities in this region. These developments remain important and will be taken into consideration, but the focus here is on filling the cultural and linguistic gaps in scholarship on Kaoka in the Philippines.

The people from south Fujian, known as the Hokkiens (*ban lam lang* 閩南人), came in large numbers and migrated to different parts of Southeast Asia.[2] Despite Southeast Asia's relatively early interactions with Chinese people

---

[1] Part of this definition is adapted from the Cambria Sinophone World Series, http://www.cambriapress.com/cambriaseries.cfm?template=85, retrieved 25 April 2020.

[2] In this chapter, I use Hokkien interchangeably with Southern Min (*ban-lam* in the Hokkien vernacular, referring to south Fujian) unless otherwise stated, as the former continues to be a term commonly used to refer to people and the language from south Fujian. However, Dory Poa, a Filipino Chinese, has alerted me that the Filipino Chinese community, particularly those whose ancestral origin was from south Fujian, often refer to their language as *ban*

---

from this region, with migration occurring as early as the sixteenth and seventeenth centuries, south Fujian is understudied particularly in its sociocultural development. Given their historical importance in Southeast Asia, it is useful to study how the language and culture of the Hokkien migrants – though originating from southern Fujian – was later subjected to the sociocultural and political contexts of the areas they migrated to, which resulted in some distinct differences across Southeast Asia.

Despite their long-time involvement and significance in Southeast Asia, studies on the language and culture of the Hokkien community in the Philippines remain limited. This chapter draws on various disciplines – including linguistics, sociology, anthropology, and religious studies – to achieve a more well-rounded understanding of the topic, especially when sources are scant. Henning Klöter's (2011) work on Early Manila Hokkien (EMH) is noteworthy, as he engaged in a pioneering study of this significant but otherwise largely ignored area. The study of EMH attests to the important role played by the Hokkien community as early as the seventeenth century. A comparative study of the Kaoka texts and EMH, as will be attempted here, can reveal some trends and changes in the development of the Hokkien vernacular in the Philippines. It furthermore underscores some local specificities of Sinophone culture,[3] by investigating how the Hokkien spoken in the Philippines differs from that of other parts of Southeast Asia and Fujian. In the 1980s, Gloria Chan Yap conducted a comprehensive study on Hokkien loanwords in Tagalog (Yap 1980). A more recent study by Wilkinson Daniel Wong Gonzales on Philippine Hybrid Hokkien (PHH) substantiates the continued importance and relevance of Hokkien among the Chinese population in the Philippines (Gonzales 2018). Yet despite these illuminating studies, there are some distinct differences between spoken Philippine Hokkien and language of the Kaoka texts. The latter forms this chapter's main focus of analysis.

Kaoka in the Philippines has existed for a century, but scholarly discussions on this theatrical form remain scant. Noting the popularity of various theatrical forms brought along by successive waves of migration from China to Southeast Asia, Chinese scholars occasionally mention Kaoka in the Philippines, but information remains scattered.[4] Despite their lack of a coherent analysis, these studies remain important as they provide a basic understanding of the way

---

lam ue 閩南話 and their people as *ban lam lang* 閩南人 rather than Hokkien. With this in mind, the term Hokkien is used with caution.

3   According to Shih's pioneering study of the Sinophone (2013: 7), "Sinophone culture is place-based and belongs to the place where it is produced".

4   Cf. Zhuang (2006), Bai (2011: 64–68), and Wu (2006).

Kaoka developed – especially before the Second World War (1942–45) – which is lacking in English scholarship. Sources about Kaoka written by Filipino Chinese, also known as Tsinoys, are equally noteworthy, as they represent their efforts and interest to document a traditional theatrical form from their home country. Anson Yu, for example, is a heritage enthusiast and has written a number of articles on Kaoka. Two of his articles were published in *Tulay*, a fortnightly Filipino Chinese newspaper still in print today (Yu 2007, 2013). There is also a study by Percy G. Ng on the socio-historical context of the genre. This author's interviews of present-day Kaoka performers have been particularly useful for my own research (Ng 2016). Josh Stenberg's recent publication on Chinese opera (戲曲 *xiqu*) in the Philippines further illustrates the relatively early existence of various forms of traditional Chinese theatre in this country, compared to the rest of the Southeast Asian region (Stenberg 2020: 58–89).

This chapter combines two main research objectives. First, it provides a succinct timeline indicating when Kaoka was transmitted to the Philippines – from its period of transmission up to the present day – as previous literature outlining this development is absent. Second, it conducts a preliminary analysis on the Kaoka playscripts, which can contribute to an understanding of the sociolinguistics of the Filipino Hokkien community, including the documentation of this theatrical form once popular across Southeast Asia but today only known in the Philippines.

2   Kaoka in the Philippines: Past and Present

Chinese migration and settlement in the Philippines was relatively early compared to counterparts elsewhere in Southeast Asia. There is evidence indicating Chinese presence prior to the arrival of the Spanish in the sixteenth century.[5] However, it is questionable to assume that Kaoka appeared in this period.[6] My conviction, that Kaoka appeared in the Philippines during the early twentieth

---

5 In his study of the language of Early Manila Hokkien (EMH), Klöter (2011: 172–73) holds that the time period of early Chinese (known as *Sangley*) settlement was in the late fifteenth century.

6 Using a rich variety of Chinese, English, and Spanish sources, Stenberg (2020) has outlined the history of Chinese opera (*xiqu*) in the Philippines from the sixteenth century. However, as emphasized in the present chapter, it is unlikely that Kaoka appeared in the Philippines before the twentieth century. As my discussion is focused on Kaoka, other theatrical forms are beyond its scope.

century, is supported by a number of observations.[7] Kaoka – pronounced as *kau-kah* in the Southern Min (閩南 *Minnan*) or Hokkien language,[8] and more commonly known in contemporary China as Gaojia opera 高甲戲 – only matured into a theatrical genre in south Fujian during the nineteenth century. Due to the high demand of the Chinese communities abroad, Kaoka opera troupes from south Fujian regularly travelled to perform in Southeast Asia, particularly from the 1840s to the 1930s. Chinese sources indicate that the earliest existence of Kaoka in the Philippines can be traced to the 1920s, particularly through the establishment of local troupes known as 'Luzon troupes' (呂宋班 *Lüsong ban*). So far, no source provides an explanation why these troupes were so named, but Luzon was undoubtedly a site of significance among the Chinese community in the Philippines.[9]

---

7   This is also supported in a paper of somewhat ambiguous authorship (*Fei hua* 2007: 802, 809), which shows that when America defeated Spain towards the end of the nineteenth century and took over the latter's control of the Philippines, their more relaxed rule led to the influx and flourishing of Chinese migrants and their hometown culture, including Kaoka. Some sources state that this paper was authored by Li Li 李麗, but the actual reference does not specify the author's name. Hence, I will refer to this source as "Fei hua 2007" in this chapter.

8   The term *Kaoka* is used throughout this chapter instead of Gaojia 高甲 opera, as the latter did not appear until 1949. Kaoka is the term used by the performers whom I have interviewed in the Philippines. It is etymologically interlinked with other Hokkien words, including *káu-kah* 九甲 ~ 九角 'nine character-roles', believed to derive from the 'seven-child' or 'seven-actor' roles in its predecessor, Liyuan opera 梨園戲. Two types of painted faces were later added to form the nine character-roles. These are the male (*sheng* 生), female (*dan* 旦), painted face (*jing* 淨), bearded male (*mo* 末), elderly female (*tie* 貼), supporting role playing miscellaneous roles (*wai* 外), and villain (*shumei* 豎眉). Another similar term is 'armour and dagger' (*gejia* 戈甲) to denote the military costumes and weapons used in Kaoka on stage, the highlight of which is often martial plays (*wuxi* 武戲). The word furthermore resembles the term 'standing on the high stage and wearing armour' (登高臺穿盔甲 *deng gaotai chuan kuijia* or 高甲 *gaojia* in short), denoting the military plays performed in Kaoka. Finally, when Kaoka spread to Southeast Asia during the nineteenth century, it won great acclaim and was regarded as 'high quality and A-grade' (高等甲等 *gaodeng jiadeng* or 高甲 *gaojia* in short) by the local Chinese community (Zhuang 2008: 136, "Fujian sheng" 2000: 33). All these folk etymologies contribute to the specific meanings the word Kaoka has acquired over time.

9   According to Hung (2014), when the Spanish colonized the Philippines in 1571 and before Koxinga 國姓爺 (Zheng Chenggong) occupied Taiwan – allowing the large influx of migrants from south Fujian – Luzon, the main island of the Philippines, appeared to be the only destination for migrants from China. Possibly also due to its proximity, most of the migrants who travelled to Luzon during this period and eventually settled there were from Zhangzhou (south Fujian).

Two associations for traditional southern Fujianese music (南音 *Nanyin*),[10] notable for their role in popularizing Kaoka to the migrant and local Chinese populations in the Philippines, were the Philippine Sy Tiok Music Association (絲竹尚義社 *Sizhu shangyishe*) established in 1922 and the Philippine Song Lim Musical Association (桑林陽春總社 *Sanglin yangchun zongshe*) established in 1923.[11] As will be discussed later, Nanyin is closely linked to Kaoka and the theatrical form is sometimes referred to as 'the (excellent) theatre of Nanyin' (南音佳劇 *Nanyin jiaju*).[12] It was also during the 1920s, arguably the peak of Kaoka in Southeast Asia, that representatives of these associations from the Philippines travelled to China and invited performers to the theatres in Manila, situated mainly at Ongpin Street (王彬街 *Wangbin jie*) and possibly Alonzo Street (阿籠計街 *Ahlongji jie*). Given its popularity with the Chinese, these performers were employed for a duration ranging from a few months to a year or two (Bai 2011: 64; *Fei hua* 2007). The Luzon troupes also mark the period when Kaoka thrived in the Philippines. Their activities came in five phases: the first Luzon troupe from 1919 to 1921, the second in 1921, the third from 1922 to 1923, the fourth from 1934 to 1936, and the fifth from 1936 to 1937.[13] The first Luzon troupe, also one of the most notable, was organized by Li Zaiju 李仔居 of the Song Lim Musical Association, who returned to south Fujian and invited the performers there to travel to the Philippines and perform at the Xintang theatre 新堂戲院 on Ongpin Street. Over the course of more than a year, more than 400 shows were staged, indicating its popularity among the Chinese in the Philippines (Bai 2011: 65).

From the 1920s to 1930s, the repertoire staged by the Kaoka performers typical of traditional theatre in south Fujian, including *Flood at Jinshan Temple*

---

10   *Nanguan* or *Nanyin* emerged during the Tang Dynasty. It was believed that during the reign of Emperor Xizong 唐僖宗 in 885 CE, the brothers Wang Chao 王潮 and Wang Shenzhi 王審知 led their army to the Min area, bringing over "big tunes" (大曲 *daqu*). *Nanyin* preserves some musical influences of the Central Plain, while assimilating local Min music (Chen & Xu 2009: 521).

11   These dates were obtained from Zhuang (2012: 764).

12   Scholars have different views on whether "Nanyin" theatre refers to Liyuan opera or Kaoka. There is also a third view, that Nanyin or Nanguan 南管 theatre refers to both theatrical forms (Shih 2012: 10).

13   The active years of the Luzon troupes slightly differ in Li's source, namely 1921–23, 1921–22, and 1923–24. Due to the lack of other sources, we can only deduce that the Luzon troupes were active around the 1920s and possibly before the outbreak of the Second World War (*Fei hua* 2007; Bai 2011: 65–66). Note also that Nanyin in the Philippines existed during the nineteenth century, so the Luzon troupes would have existed before the establishment of some of the Nanyin associations. Further detail is beyond the scope of this chapter and will require more research.

(水淹金山寺 *Shuiyan Jinshansi*), *Ziya Slays Three Demons* (子牙捉三妖 *Ziya zhuo sanyao*), *Datong City* (大同城 *Datong cheng*), *Disguising as Elderly Couple* (公婆拖 *Gongpo tuo*), *The Jade Lovebirds Fan* (玉骨鴛鴦扇 *Yugu yuanyang shan*), *Rousing the Bridal Chamber* (鬧洞房 *Nao dongfang*), *The Foreign Woman Causing a Commotion* (番婆弄 *Fanpo nong*), *Tang Er Leaves his Wife* (唐二別妻 *Tang Er bieqi*), *Beating the Wife* (姈婆打 *Lingpo da*), *The Frantic Monk Scolds Qin Kuai* (瘋僧罵秦檜 *Feng seng ma Qin Kuai*), and *Guan Fu Leaves his lover* (管甫送 *Guan Fu song*).[14] A number of these shows were popular stories emphasizing the clown (丑 *chou*) role, which was a common feature of Kaoka.[15]

The performance of Kaoka came to an abrupt halt with the outbreak of the Second World War and the Japanese occupation of the Philippines. Kaoka performances that catered for traditional celebrations, such as weddings, birthdays for elders, and religious festivities, resumed after the war (Bai 2011: 65). While it would appear that the cultural and entertainment scene was similar to the pre-War situation, the socio-political climate – most notably after the Philippines' declaration of independence in 1946 and the outbreak of the Cold War in 1947 – had a drastic impact on international and familial relations. This in turn affected Kaoka. Taiwan, instead of mainland China, soon became a close ally of the Philippines, especially because of geopolitical and ideological concerns. The post-War era also saw the increasing reliance of both the Republic of China (ROC) in Taiwan and of the Philippines on the United States.[16] Their alliance was further heightened by the containment strategy adopted by the United States during the Cold War. Sharing a common ancestral origin with the Taiwanese, the Chinese in the Philippines, particularly those of south Fujian origins, also began to demonstrate closer cultural affiliations with the former.

As mentioned earlier, the Nanyin associations played a significant role in bringing Kaoka performances in the Philippines to their peak during the 1920s. Although the political climate of the post-War period went through a dramatic transition, the role of Nanyin and its theatrical forms continued to be – to quote

---

14 This repertoire is given in Bai (2011: 65–66).

15 As Gaojia opera began to mature as an opera form complete with makeup, music, singing, stage movements, and storyline, it also developed the distinctive trait of featuring the clown character, to the extent that this character role became synonymous with the operatic form. The significance of the clown role in Gaojia opera is mentioned in Bai (2011: 66–67).

16 The policy of "getting close to Taiwan and away from mainland China" (親台灣遠大陸 *qin Taiwan yuan Dalu*) as well as the strong influence of the United States on the Philippines is mentioned in *Fei hua* (2007: 825) and Wickberg (2006: 22).

Tan and Rao in their definition of Sino-soundscapes – "intentionally produced and created in ways to register distinctive identity and articulate cultural positions in particular places" (Tan & Rao 2016: 6). This form of cultural exchange between Taiwan and the Philippines was evident.

During the Cold War era, Taiwanese troupes allowed to travel overseas were often backed by a political agenda, exemplified by the slogan "repel the Communists and resist the Soviets" (反共抗俄 *fan gong kang e*). In 1958, the Hsin Li-yuan troupe 新麗園劇團 – formed by the Pao-tao (寶島; lit. 'Treasure island', a popular reference for Taiwan) Gezai opera troupe and the exclusive Gaojia opera troupe for the military (軍中專屬高甲戲團 *junzhong zhuanshu Gaojiaxi tuan*), led by the Airforce leader Lin Chin-chih 林金池 – visited the Philippines in the name of establishing a rapport with the local Chinese. The visit lasted for about three months (*Fei hua* 2007: 828). Later, in August 1958, the I-chun Yuan troupe 宜春園, established by the I-chun Yuan Gezai opera troupe and Hsin chin-chu 新錦珠 Gaojia opera of Taichung, was invited by a Filipino Chinese man named Gao Qingyun 高慶雲. It should be noted that a number of Taiwanese troupes that visited the Philippines were joint Gezai-Gaojia ensembles. The name of the troupe usually followed that of the Gezai opera troupe, which suggests that Gezai opera was the key player in these performances, whereas Gaojia opera performances only played a supporting (客串 *kechuan*) role. This is not surprising in view of the general trend of Gezai dominance over the older Gaojia opera in Taiwan as well as other parts of the Southeast Asia, such as Malaya and Singapore. However, this was not the case in the Philippines. According to the recollection of Gaojia opera performer Chen Hsiu-Feng 陳秀鳳, "[t]o the surprise of the troupe, the supporting role played by Gaojia opera stole the limelight from Gezai opera. As a result, Gaojia opera performances were continuously added and when the troupe released their program, the tickets were sold out three days before the performance" (*Fei hua* 2007: 828). The prevalence of Kaoka, despite its dwindling role in other parts of Southeast Asia and Taiwan where the Hokkien communities were most populous, continued to enjoy strong support from the Chinese community in the Philippines. I would argue that this persistence was closely associated with the unifying vigour of the Nanyin tradition, yielding a significant Sino-soundscape rooted in the nineteenth century.[17]

---

17   For example, two Nanyin associations, namely Changhe Langjun she 長和郎君社 and Jin Lan Langjun she 金蘭郎君社 were said to be established during the early nineteenth century (*Fei hua* 2007: 807).

3   **Local Kaoka Establishments**

Socio-political developments after the Second World War did not dampen the enthusiasm of the Filipino Chinese for Kaoka. The Philippines became one of the major centres in Southeast Asia to organize Kaoka-related activities. Another factor that prompted local developments and establishments of Kaoka troupes was the inability of many performers and musicians to return to their hometown due to the civil war between the Nationalists and the Communists and the resulting political instability in mainland China. Since there was a strong support for Kaoka, they decided to settle or temporarily remain in the Philippines (*Fei hua* 2007: 842). In 1950, the Shengxing troupe 勝興劇團 was established. According to Teresita Bee Hua Chan 曾美華 (hereafter referred to as Teresita Chan), her teachers Wu Yumu 吳于目 and Wang Shanda 王善達, who were natives from Amoy, decided to remain in the Philippines when the war broke out and eventually established their troupe when the war was over. In 1951, Nanguo Drama Society 南國劇藝社 was founded. The Xiu Lian Xing troupe 秀聯興劇團 was established in 1962,[18] and the Kim Siu Eng troupe 金秀英劇團 in 1965 by Kim Siu Eng, a renowned performer and foster mother of Teresita Chan.[19] Particularly in the case of the Nanguo Drama Society and the Shengxing troupe, which made efforts to transmit the theatrical form, children were recruited as apprentices of Kaoka and supervised by experienced teachers. These local troupes often performed in traditional contexts, particularly those related to deities' feast days and celebrations organized by various associations. For example, on 20 June 1957, the Song Lim Musical Association invited the Shengxing troupe to perform in celebration of the feast day of Master Guan (Guan Fuzi 關夫子) and on 10 April 1963, the Dadao Shrine 大道玄壇 invited the Xiu Lian Xing troupe to perform on the feast day of their Grand Master Yun Meng 雲夢祖師 (*Fei hua* 2007: 842–43). These local establishments were able to continue the legacy in collaboration with their predecessors from Fujian, but their success was met with various challenges ahead.

The 1970s was a watershed for Kaoka in the Philippines, which saw a change in the enthusiasm and support for this form of traditional theatre. Locally established troupes performed in the traditional context, that is, an outdoor stage erected near or within the temple compound.[20] As mentioned

---

18   Pan (2013: 222), cited in Stenberg (2020: 77).

19   *Fei hua* (2007: 842). Personal interview, Teresita Bee Hua Chan, 6 October 2017. For troupes where the actual spelling is not known, the *hanyu pinyin* version is used instead.

20   Unlike their predecessors, locally established troupes were said to perform in temple contexts and less so in indoor theatres (*Fei hua* 2007: 842). As recalled by Hau (2020) cited in Stenberg (2020: 72), makeshift stages performing Kaoka used to be erected on the streets.

FIGURE 6.1 Kaoka performance inside Qing Long Dian without any stage setup
OWN WORK

earlier, local troupes performed on the feast days of deities during the 1950s
and 1960s. These outdoor performances came to a halt when then President
Ferdinand E. Marcos issued Proclamation No. 1081 on 21 September 1972,
declaring martial law in the Philippines.[21] The martial law period (1972–81)
meant that curfews were enforced throughout the country that lasted from
12 midnight to 4 in the morning. Those found loitering on the streets at these
hours were detained. The curfew affected Kaoka performances, which could
previously go on for hours till near midnight. For fear of being detained or
questioned, the patrons of Kaoka diminished. With a sharp drop in audiences
and supporters, temple invitations of Kaoka troupes were drastically affected
too. For example, a seven-day performance was reduced to three days, which
greatly impacted on the livelihood of troupes (Ng 2016: 57). Today in Binondo
(known as Manila's Chinatown), Kaoka performances are staged inside the
temple premises without any stage setup (Fig. 6.1).

The Hoc Kian troupe led by Teresita Chan has been selected for this chapter,
particularly because the troupe leader was associated with older local estab-
lishments. These include the founders of the Shengxing troupe, who were
her teachers, and Kim Siu Eng of the Kim Siu Eng troupe, who was a personal
teacher and foster mother of Teresita Chan. During my fieldwork conducted in
Manila and Bulacan in 2017, the Hoc Kian troupe appeared to be quite active

---

21 Ng (2016: 56); Official Gazette, Declaration of Martial Law, retrieved 4 December 2019:
    https://www.officialgazette.gov.ph/featured/declaration-of-martial-law/.

and within a week, I observed two performances respectively at Si Wang Fu (鎮池宮古坑四王府 *Zhenchigong gukeng siwangfu*) on 1 and 2 December, and Qing Long Dian 青龍殿 on 3 and 4 December.

From my observation on the Hoc Kian troupe led by Teresita Chan, it appears that she conforms to the traditional context as much as possible in terms of the music, repertoire, and performance. All these aspects contribute to the sustenance of the Sino-soundscape as expected by the larger traditional Filipino Chinese community. This sustained Sinophone heritage is also evident in the textual representation of Kaoka. Teresita Chan revealed that her proficiency in Chinese, whether spoken or written, was low, as she did not pay a lot of attention to learning the language during her school days. However, the knowledge or craft literacy required for Kaoka meant that she had to learn various aspects of the Sinophone: knowing the Chinese characters (Sinographs) of the repertoire and creating romanizations of the Hokkien vernacular, in this case the Kaoka playscripts, so that performers could learn them.[22]

The repertoire, as presented in Table 6.1, is an outline of various story titles (Fig. 6.2) written by Teresita Chan by hand, including in a non-standard orthography shown in the leftmost column. For story titles that contain non-standard orthographic characters, I have included the corresponding standardized characters to provide the reader with a better sense of the story or main character in relation to the traditional Hokkien theatre or other Chinese opera forms. For example, *Wang Lihua Descends from the Mountains* (王麗華下山 *Wang Lihua xia shan*) actually refers to *Fan Lihua Descends from the Mountains* (樊梨花下山 *Fan Lihua xia shan*), which will be analysed later. Teresita explained that it is required of her as the troupe leader to write this repertoire (Fig. 6.2), as the temple personnel will usually seek permission of the deity through the tossing of divination blocks (Hokkien: *puah-pue* 跋桮; Mandarin: *zhijiao* 擲筊) to determine which story title is to be performed. This traditional way of selecting the story title from a repertoire is also observed in the Hokkien theatre of Quanzhou.[23]

---

22   Craft literacy is regarded as special knowledge usually limited to a specialized group engaged in the same craft. For a more in-depth discussion on this, see Goody (1975).

23   As Ruizendaal (2006: 152) observes in his study of the marionette theatre in Quanzhou, "the names of the plays were written on a traditional harmonica-like list … the list was opened at the name of one play and placed on the altar and by the throw of the divination blocks the deity would decide which play(s) was to be performed".
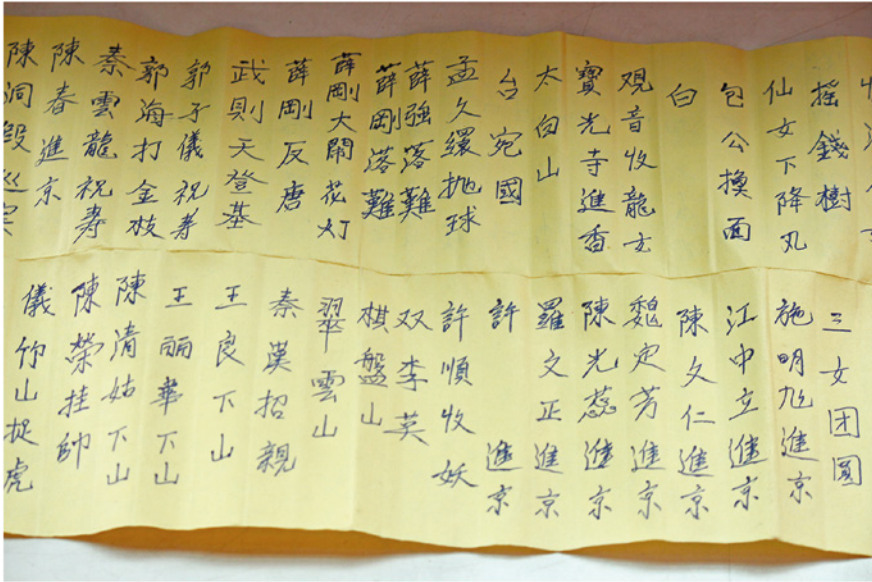
FIGURE 6.2 An image of part of the repertoire written by Teresita Chan
OWN WORK

TABLE 6.1   Repertoire of the Hoc Kian troupe

Repertoire as written by Teresita Chan	Default title	English translation
八仙過海		Eight immortals cross the sea
王母祝壽		Wishes on Queen Mother's birthday
李洞彬收妖	呂洞賓收妖	Lü Dongbin subdues the demon
四仙記		Story of the Four Immortals
收河仙姑	收何仙姑	Subduing He Xiangu
搖錢樹		The coin-shedding tree
仙女下降凡		The fairy descending to the mortal world
包公換面		The real and fake Justice Bao[a]
觀音收龍女		Guanyin accepts Dragon Girl as disciple
寶光寺進香		Offering incense at Baoguang temple

a   My speculation is that this is related to the story *The Real and Fake Justice Bao* 真假包公. However, this will require further verification based on the details of this story.

TABLE 6.1   Repertoire of the Hoc Kian troupe (*cont.*)

Repertoire as written by Teresita Chan	Default title	English translation
太白山		Mount Taibai
台宛國		Kingdom of Tai Wan
孟久繯抛球	孟九環抛球	Meng Jiuhuan tosses the bouquet[b]
薛強落難		The misfortune of Xue Qiang
薛剛落難		The misfortune of Xue Gang
薛剛反唐		Xue Gang goes against Tang kingdom
武則天登基		Wu Zetian ascends the throne
郭子儀祝壽		Guo Ziyi offers his birthday wishes
郭海打金枝		Guo Hai beats Jin Zhi
秦雲龍祝壽	秦雲龍祝壽	Qin Yunlong offers his birthday wishes
陳春進京		Chen Chun travels to the capital
陳洞假巡案		The Fake Investigation by Chen Dong
李淵登基		Li Yuan ascends the throne
王寶釧抛球		Wang Baochuan tosses the bouquet
湘江會		The meeting at Xiang River
麒麟山		Mount Qilin
楊宗保取木棍		Yang Zongbao retrieves the wooden stick
穆桂英招親		Mu Guiying invites the groom
征南戕國	征南戰國	Heading south for war
三女團圓		Reunion of the three maidens
施明乣進京	施明旭進京	Shi Mingxu travels to the capital
江中立進京		Jiang Zhongli travels to the capital
陳夊仁進京[c]		Chen Jiuren travels to the capital
魏定芳進京		Wei Dingfang travels to the capital
陳光蕊進京		Chen Guangrui travels to the capital
羅文正進京		Luo Wenzheng travels to the capital
許？進京[d]		Xu travels to the capital
許順收妖		Xu Shun subdues the demon

b   This story is associated with *Xue Gang Goes Against the Tang Dynasty*, also listed in the repertoire.

c   It is unclear whether this story refers to 陳久仁 or 陳九仁.

d   Only the last name is provided, as opposed to the full name.

TABLE 6.1   Repertoire of the Hoc Kian troupe (*cont.*)

Repertoire as written by Teresita Chan	Default title	English translation
双李英	双李英	e
棋盤山		Mount Qipan
翠雲山		Mount Cui Yun
泰漢招親	秦漢招親	Qin Han marries into the Bride's family
王良下山		Wang Liang descends from the mountain
陳清姑下山	陳靖姑下山	Chen Jinggu descends from the mountain
王麗華下山	樊梨花下山	Fan Lihua descends from the mountain
陳榮挂帥		Chen Rong assumes command
儀竹山捉虎		Hunting for the tiger at Mount Yizhu
蘇后取頭闵f	蘇后取頭關	
借子龍退典禮		Using Zilong to depart the ceremony
杜世光比武		Du Shiguang displays his martial skills
岳飛收楊再兴		Yue Fei defeats Yang Zaixing
取潼闵	取潼關	Invasion at Gate of Tong
取边闵	取邊關	Invasion at the borders
白鹤闵	白鶴關	At the Gate of Bai He (White Crane)
西龍闵	西龍關	At the Gate of Xilong (Dragon of the West)

e This is also a story that requires further verification, as I am unsure whether the name Li Ying refers to the main character.

f I have not been able to identify a story related to Su Hou, so the English translation is not provided here.

Non-standard orthography is common in handwritten scripts and I attempt to accurately represent these original characters (夂, 秦, 闵, 边, 英, 尢, 戝). If the story title conforms to the standard orthography and/or a story usually known in traditional theatre, the part in the middle column is shaded in grey.

As observed in this repertoire, themes that frequently appear include travelling to the Capital, invasions, and descending from mountains. The phrase "travelling to the Capital" can suggest the emphasis on scholarly pursuit often associated with civil plays (*wenxi* 文戲), whereas "invasions" and "descends from the Mountain" suggest, respectively, martial plays (*wuxi* 武戲) and stories related to divine arts.

4   A Preliminary Analysis of the Kaoka Playscripts

To my knowledge, the Kaoka playscripts created by Teresita Chan, the current troupe leader of the Hoc Kian troupe in the Philippines, have not received any academic attention. Teresita Chan saw the need to write these playscripts so that new performers, especially those who started with little or no proficiency in Hokkien, can learn the lines. She mentioned to me that her teachers taught her the lines orally, so there was no text prior to her creation. These playscripts attest to the existence of a particular Sino-soundscape as mentioned earlier, notably in Southeast Asia and beyond (Taiwan and Kinmen in particular).[24] The Philippines is currently the only location within this Sino-soundscape that has sustained the theatrical form closely associated with Nanyin.[25] While Nanyin itself is still performed in other sites in this Sino-soundscape, Kaoka has not been able to "withstand" the popularity of Gezai opera, its more modern counterpart from Taiwan, as the latter adopts a musical style distinct from Nanyin. The fondness for Nanyin can be regarded as a marker of Filipino Chinese identity, as I argue in this chapter, which has allowed Kaoka to survive in the Philippines. As Kim Chew Ng noted, "a writer in China or Taiwan is perhaps preordained to write in the Sinitic script. Abroad, especially in a country where Hua people are oppressed and regarded as a minority, a writer's choice of the Sinitic script involves a serious value judgement and has significance in terms of cultural identity … th[is] decision … requires an exertion of effort, as the language is something that must be 'acquired'" (Ng 2013: 79). The Kaoka playscripts also represent the consistent effort by practitioners to sustain this tradition, even if many do not have the proficiency to write Chinese characters. For the Hoc Kian opera troupe, I have observed that older performers are of Chinese descent, that is, one of their parents or both are Chinese, whereas the younger performers are Filipinos with no Chinese descent. Some of the older performers communicated with me in Mandarin, while Teresita Chan, the troupe leader, communicated in Hokkien and some English. The Kaoka playscripts (Fig. 6.3) are essential learning materials for the performers,

---

24   Quite similar to the Philippines, Gaojia opera (Kaoka) was the more dominant theatrical form in Kinmen, compared to Gezai opera. The most recent record of a local Gaojia opera troupe was in 2000, but more research is required to understand the situation in Kinmen. For a contextual understanding, see Chia (2019: 76–79).

25   Stenberg (2020: 71–72) also notes that Nanyin (also known as Nanguan) musicians in the Philippines were pivotal in sustaining a related theatrical form in Taiwan, known as *nanguan xi* 南管戲.
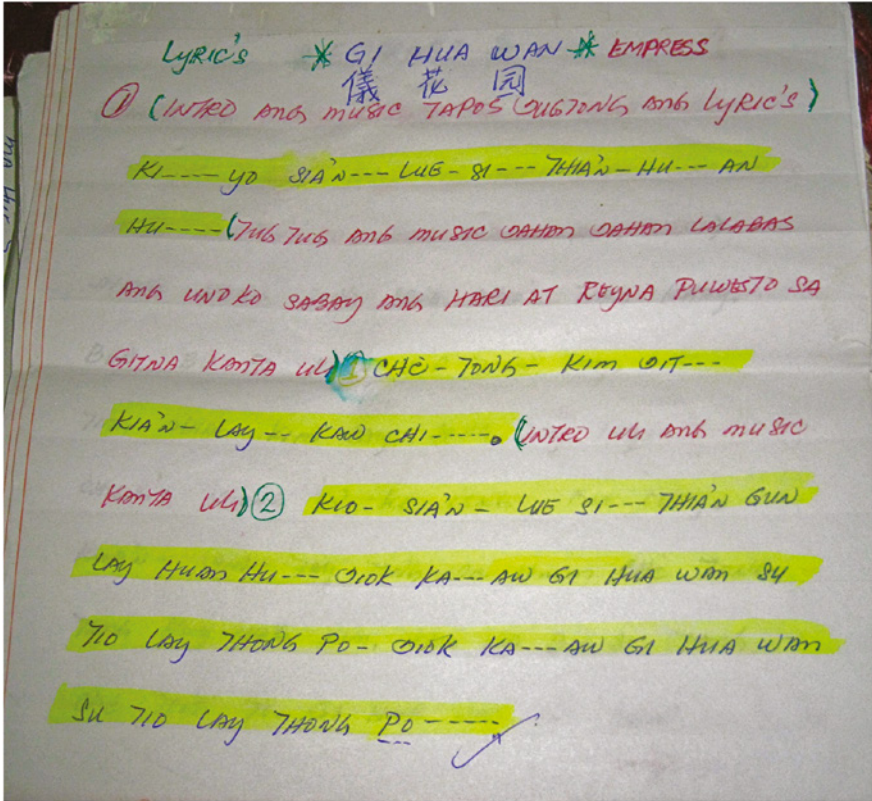
FIGURE 6.3 An image of a Kaoka playscript by the Hoc Kian troupe, photo taken by author with permission of the troupe
OWN WORK

especially the younger generation who often start out without any proficiency in Hokkien.

Taking a closer look at the playscripts, they are written in a style quite different from the aforementioned studies on colloquial Hokkien varieties in the Philippines. Compared to studies on the more contemporary version of Hokkien, the Kaoka texts reflect less everyday usage and more theatrical language, in which the performers were specialized. In analysing the selected playscripts, I have also discovered the occurrence of stock phrases commonly used among Hokkien performers outside the Philippines, which will be further discussed below.

Studying these Kaoka playscripts has been challenging, as there are a number of linguistic skills required to provide an informed analysis. First, not being

a native from the Philippines meant that I was unaware of how Tagalog has influenced the romanized Hokkien script in use. To overcome this lack of proficiency, I consulted with Filipino native speakers who were not ethnically Chinese or Chinese-descended, which meant that they did not speak Hokkien (*Ban-lam-ue*). They provided some insight about the explanatory notes written in Tagalog, noting that some of the words were misspelt, possibly due to the limited literacy of the scribe or the fact that the script was an earlier form of Tagalog – although the precise time period of the script is unclear – that is less common today.[26] Secondly, knowing Hokkien is useful to analyse the texts, but this is particularly the case when researching colloquial Philippine Hokkien varieties. The Kaoka playscripts additionally require knowledge of the theatrical language, particularly the one used in Hokkien theatre, in which I have acquired proficiency given my background in this field. To help me with this chapter's preliminary analysis, I have also consulted with different people proficient in both Hokkien and Tagalog, and used various Quanzhou, Zhangzhou and Taiwanese dictionaries as references.[27]

The theatrical language in romanized Hokkien is an important characteristic of the Kaoka playscripts. Table 6.2 illustrates the corresponding Chinese characters that are pronounced (and romanized) differently in the literary and colloquial styles. This reflects a broader phenomenon in the Hokkien language (see also Ding and Churchman in this volume), which is said to be one of the most complex among the Chinese regional languages in terms of differences in colloquial and literary reading (*wenbai yidu* 文白異讀).[28] For the sake of comparison, the rightmost column includes Klöter's record of EMH corresponding to the same Chinese character.

---

26   I am grateful to Dr. Darlene Machell De Leon Espena from the Singapore Management University (SMU), who helped me with understanding the Tagalog translation in the Kaoka scripts. The topic of observed differences between the Tagalog translation and present-day Tagalog usage, which requires the expertise of Tagalog linguists, falls beyond the scope of this chapter. As the Tagalog sections provide explanatory notes, a general understanding sufficed to help me analyse the source text.

27   The dictionaries used in this analysis include Zhou (2006) and *Taiwan Minnanyu Changyongci cidian*.

28   Lin (2008: 49–51). Although Lin suggests that the most complex variety – in terms of such differences in literary and colloquial reading – is the Quanzhou vernacular, I think it is hard to measure its complexity. The coexistence of colloquial and literary readings applies to the Zhangzhou and Taiwanese vernaculars as well, although I will limit the discussion to the Kaoka playscripts.

TABLE 6.2   Literary Hokkien in Kaoka playscripts

Literary Hokkien of Kaoka playscripts	Corresponding Chinese characters	Colloquial Hokkien	Early Manila Hokkien (Klöter 2011)	English gloss
bio	母	*bu*		mother
chian	前	*tsing*	*chēng, cing⁵*	front
din (*lin*)	人	*lang*	*lang, lang⁵, lāng*[a]	person
ha	下	*e*	*ě, ee³*	to descend
hak	學	*oh*		learn
hua	花	*hue*	*hua, hue*	flower
hun	分	*pun*	*hún, hun¹*	separate
ngo	我	*gua*	*goa, gua*[b]	I
san	山	*suann, sua*	*san¹, suă*	mountain
ti	知	*tsai, zai*	*chai*	to know
to	到	*kau*	*câu, kau⁷*	to reach

a   Recorded by Gonzales (2018: 33) as *láng* and by Yap (1980: 110, 132) as *lang, láng*, and *láŋ*.
b   Gonzales (2018: 34) records it as *guâ*.

Performers of traditional Hokkien theatre use a mix of literary and colloquial registers. For example, characters involving learned men, officials, and the royal court speak and sing more in the literary style, whereas those of a lower status, such as the commoners, use a more colloquial form. Performers tend to acquire or learn phrases in literary style as stock phrases, so that they can be applied whenever the need arises. Some examples of these stock phrases in the literary style also exist in the Kaoka playscripts, as listed in Table 6.3:

TABLE 6.3   Stock phrases of literary Hokkien observed in the Kaoka playscripts

Stock phrases	Kaoka playscripts	Corresponding English meaning
免禮	*bian le*	dispense with ceremony
且慢	*chia ban*	please wait
妾身	*chiap sin*[a]	wife (referring to oneself)

a   The character 身 'body' is recorded by Klöter (2011: 326, 327) as *sin* and *syn*.

TABLE 6.3   Stock phrases of literary Hokkien observed in the Kaoka playscripts (*cont.*)

Stock phrases	Kaoka playscripts	Corresponding English meaning
膝下無嗣	*han ha bu si*	no progeny
奉萬歲聖旨	*hong ban sue sing chi*	following orders of the Emperor
飲酒作樂	*im chiw chok lok*	drink and make merry
共守共結髮	*kang siw kang kiat huat*	to stay together in marriage
(夫妻) 結髮	*kiat huat*	marriage of husband and wife
叫聲老爺	*kio sia'n law ya*	I call to you, my lord
保庇我夫早轉/回來	*po pi gua hu cha teng lay*	bless my husband for an early return

The stock phrases listed in Table 6.3 are found in the Kaoka playscripts and display similarities with the theatrical language that I have studied elsewhere (Chia 2018: 31, Table 6.3). Such similarities demonstrate that Kaoka performers in the Philippines adhere to the language and performance conventions observed in Hokkien theatre more generally.

There are about 35 to 40 Kaoka playscripts that I was given permission to read and document. As it is not possible to cover all within this chapter, I will provide an overview of the playscripts. A number of the texts include translations into Tagalog and appear to be general lines for performers. The *Gi Hua Wan* (儀花園 *Yi Hua Yuan*) 'Righteous Garden', for example, features the Emperor's and Empress' lines in romanized Hokkien and Tagalog (see Fig. 6.3 above). Flipping through these playscripts, a recurring theme is the imperial court. There are also song lyrics related to this theme, beckoning the question why it appeared so frequently. I linked this to my previous study of Gezai opera in Singapore. Unlike some other theatrical forms such as Teochew opera and Cantonese opera, Gezai opera generally does not have a playscript and performers rely on improvisation and learning orally, depending on their level of literacy. Despite its improvisational nature, performers related to me that they tend to learn by rote certain spoken and sung parts that are specific and allude more to the literary style, particularly for roles like the Emperor. This is harder to improvise and so they would learn through written texts, notes, or oral means and remember them by heart (Chia 2018: 28). The need to remember lines in the literary style by rote, including those of the Emperor, could explain why they frequently appear in the Kaoka playscripts.

For the purposes of this chapter, I have selected one playscript for detailed analysis, titled *Ong Le Hua Ha Shan* (王麗華下山 *Wang Lihua xia shan*) 'Ong

Le Hua descends from the mountains'. At first glance, Ong Le Hua does not appear to be a familiar character in the Hokkien repertoire. However, through a repeated analysis of the storyline and with verification from the author Teresita Chan, I could make sense of the story. It is related to the classical character Fan Lihua (樊梨花; Hokkien: *Huan Le Hua*), since there was an episode where she was sent by her teacher Venerable Mother of Mount Li (黎山聖母 *Lishan Shengmu*, some sources have 梨山老母 ~ 驪山老母 *Lishan Laomu*) to descend from the mountains. The names of the two main characters are Ong Le Hua and her father Ong Che Sieng 王志賢 – corresponding to Huan Le Hua and Huan Hong 樊洪 – and the time period is set in the Ming Dynasty (1368–1644 CE) instead of Tang (618–907 CE). Nevertheless, the storyline remains more or less the same, possibly due to modifications in the process of transmission. Ong Le Hua is sent to the mountains at an early age to learn martial arts from the Venerable Mother of Mount Li. The latter decided that it was time she descended from the mountains to save Ong Le Hua's father, Ong Che Sieng, who was captured by foreign invaders. In this story, Venerable Mother of Mount Li's name remains the same and is key to substantiate that the Kaoka playscript is a modified version of the original story of *Fan Lihua Descends from the Mountains* (樊梨花下山 *Fan Lihua xia shan*).

To investigate the nature of the Kaoka playscripts, I first observed whether the scribe has a personal style of writing the romanized Hokkien and whether the same word is romanized consistently. From my initial observation, I noticed that words ending in ⟨*ay*⟩ correspond to /*ai*/ in the default Hokkien romanization. Examples include *lay* (*lai* 來), *tay* (*tai* 帶), *thay* (*thai* 待) and *kay* (*kai* 該). This can be observed in several other examples, which are listed in Table 6.4 below. In Yap's study of the Hokkien borrowings in Tagalog, such as *kintsay* (*khinchai* 芹菜) and *petsay* (*peqchai* 白菜), this minor orthographic difference is also attested (Yap 1980: 29–31). It is presumably the result of influence from written Tagalog, in which ⟨ay⟩ likewise corresponds to /ai/ (rather than /e/). Along similar lines, ⟨aw⟩ in the Kaoka playscripts corresponds to /au/ in mainstream Hokkien.

Table 6.5 includes the Kaoka text of *Ong Le Hua Descends from the Mountains*, along with the default Hokkien romanization in the second column. The purpose of juxtaposing both romanizations is to illustrate that the Kaoka playscript adheres quite closely to a general Hokkien romanization. As mentioned before, the Hokkien romanization provided here is informed by Quanzhou, Xiamen, Zhangzhou, and Taiwanese dictionaries. Where these varieties differ internally, I indicate those differences in footnotes or with a slash (/). I have verified this playscript with Teresita Chan, but there are challenges, as she does not write much Chinese and the script also has a number of versions, including

TABLE 6.4   Romanizations in the Ong Le Hua playscript

Chinese character	Romanization in Kaoka playscript Ong Le Hua	Generic Hokkien romanization	English gloss
你	*di*[a]	*li*	you
了	*diaw*	*liau*	already
交	*kaw*	*kau*	submit
該	*kay*	*kai*	should
快	*kuay*	*khuai*	to hasten
來	*lay*[b]	*lai*	to come
乃	*nay*	*nai*	to be
帶	*tay*	*tai*	to bring
待	*tay*	*tai*	to wait
潮	*tiaw*	*tiau*	tide

a   This is also recorded by Gonzales (2018) as *di* (a number of examples cited).
b   The character 來 'coming' is recorded by Klöter (2011: 222, 223) as *lāy* or *lai⁵*.

TABLE 6.5   Analysis of the Ong Le Hua Ha Shan playscript with Chinese and English translations

Kaoka playscript	Generic Hokkien romanization	Chinese translation	English translation
Ngo nay le san sieng bio na in ong le hua	*Ngo nai le san sing bio na in ong le hua*	我乃黎山聖母那因王麗華	I am Venerable Mother of Mount Li. Because for Ong Le Hua,
I gun si to hun chiw'	*I gun si to hun **tshiu***	伊阮是到分手（的時候）	She and we have to go separate ways
gun kay thio ha san	*gun kai tioh ha san*	阮該著下山	We have to descend from the mountains
tay tha'n to siang san, chian lay hak ge	*tai **thann** to siang san tsian lai hak ge*	帶她到雙山前來學藝	(Recalling the old times) Bringing her to the Shuang mountains to learn the art
che tay chia'n ba chit pin it si sim hiat lay tiaw	*tse tai tshiann pa tsit pinn tsit si sim hiat lai tiau*	坐待且罷這一時心血來潮	as I cannot sit there and wait. On my end, it was a decision made with impulse

TABLE 6.5 Analysis of the Ong Le Hua Ha Shan playscript with Chinese and English (*cont.*)

Kaoka playscript	Generic Hokkien romanization	Chinese translation	English translation
kay thio kut chuy'n i suan pian ti[a] *haw se se diaw … ay ya chia'n ban* *na in ong che sieng*	*kai tioh kut tshui it suan pian ti* *xau*[b]*/ho sei si liau … ai ia tshia ban* *na in ong tse/tsi tsheng*	該著骨推（揦指）一算便知 好是[c]是了 哎呀且慢 那因王志成	By pinching my fingers, I can predict This is it. Alas, wait a moment Because Ong Che Sieng (Wang Zhi Cheng)
pi huan ban tiong diaw it siong *gua kay thio dieng to te* *ha san **kue** kiw di*	*pi huan*[d] *ban tiong liau it siong* *gua kai tsu/teo ting to te* *ha san kue (khi) kiu li*	被番蠻中了一傷 我該著/得叮（囑）徒弟 下山過（去）救你	is injured by the foreign invaders I have to instruct my disciple to descend from the mountains and save you
lay bin to te ke lay	*lai bin to te ko/kua/ ke lai*	內面徒弟過來	Come inside, my disciple
put iong -o- to te lay[e]	*put iong to te lai*	不用 徒弟來	Don't have to (stand on ceremony), come on in, my disciple
na si di thia'n pi huan ban tiong diaw it siong	*na si li tia pi huan ban tiong liau it siong*	那是你爹被番蠻中了一傷	Your father has been injured by foreign invaders

a   *I* and *it* 'one' are used here, which correspond to the Quanzhou and Zhangzhou romanization. See Zhou (2006: 1071). The word *i* here should not be confused with 伊 's/he', exemplifying once again the challenges of attempting to analyse Sinophone texts that lack characters, as the same romanization can refer to different corresponding Chinese characters.

b   As indicated in Zhou (2006), *xau* can be written and pronounced as *hau*.

c   This is partly based on script with Chinese characters provided by Teresita Chan. The character 好 'good' is recorded by Klöter (2011) as *ho* in general (many examples cited throughout the book).

d   I could not locate the romanization for 番 in Zhou (2006). It is recorded by Yap (1980: 132) as *huân*.

e   In the Kaoka script, *-o-* possibly indicates a pause.

Kaoka playscript	Generic Hokkien romanization	Chinese translation	English translation
to te lay, gua dieng di ha san kue kiw	*to te lai, gua ting li ha san kue (khi) kiu*	徒弟來，我叮你下山過（去）救	Come my disciple, I instruct you to descend from the mountains and do the rescue
di thia'n din chit ke hue tit tuan uan	*li tia lin tsit ke hue tit thuan uan*	你爹恁（你们）一家回得/去團圓	Then you and your father can return home and reunite
to te la tia gua kaw tay[f]	*to te lai thiann gua kau tai*	徒弟來聽我交代	Listen to my instructions, my disciple
to te lay, tia gua chua thieng dieng[g]	*to te lai, thiann gua tshua thing ting*	徒弟來，聽我（的話）帶（這盞）廳燈	Come my disciple, bring this hall lantern
sok sok ha san ti -o-o- kuay ke	*sok sok ha san tit khuai ke*	速速下山得快計	The brief plan is to quickly descend from the mountains.
ho diaw to te khi sin	*hau liau to te khi sin*	好了徒弟起身	Alright, it is time to get moving, my disciple
gua kay thio tan thay siaw sit pa ... diaw	*gua kai tioh tan thai siau sit pa liau*	我該著等待消息罷了	I will wait here for your news

f   *La* possibly means *lay* with the omission of ⟨y⟩.
g   The character 燈 'lantern, lamp' is recorded by Yap (1980: 132) as *tiêŋ*.
There are still minor parts (bold-italicized) that are unclear to me and any error is my own.

the one she remembers and one copied by her student and current Kaoka performer Raquel Espena. I have used the latter playscript because Teresita indicated that she currently does not have the physical copy of the former and the photographs she provided were too blurry to decipher. Although versions differ, the essential content – particularly the role of Venerable Mother of Mount Li, the way she predicted that Ong Le Hua's father would be trapped by foreign invaders, and her advice to Ong Le Hua to descend from the mountains to save her father – has helped me to decipher the text. However, it should be noted that without the context it is difficult to determine the meaning and corresponding Chinese characters, especially when the romanized script has

no tonal indication, such that one romanized word can correspond to many characters of the same pronunciation.

5 Concluding Remarks

This study of Kaoka, a traditional theatrical form originating from south Fujian and still performed in the Philippines today, illustrates the diversity and complexity of the Sinophone. Its two prisms of analysis – cultural and linguistic – are deeply intertwined. The timeline I have attempted to reconstruct of the development of Kaoka in the Philippines highlights the prevalence of Nanyin as the bedrock of the interconnected Sino-soundscape that developed among the various Hokkien communities in Southeast Asia and beyond, including Taiwan and Kinmen. It was Nanyin, as I have attempted to demonstrate, that was crucial to sustain the popularity, growth, and continued existence of Kaoka in the Philippines, as the latter has disappeared from other parts of Southeast Asia. The various features of the Kaoka playscripts discussed in the second part, including the colloquial and literary readings, illuminate the linguistic implications of this story. The colloquial-literary dichotomy is less observed in previous studies on Philippine Hokkien, partly because the literary forms are less frequent in everyday usage. However, as I have shown, it is nevertheless important to obtain a more textured sociohistorical picture in order to understand Kaoka and Hokkien theatre in general. In the Philippine context, language and performance have clearly reinvigorated each other. This analysis has furthermore illustrated the preservation and continued use of stock phrases commonly seen in Hokkien theatre elsewhere. The romanization of the Kaoka playscripts closely resembles the romanization of Hokkien in Quanzhou, Zhangzhou and Taiwan, conveying a sense that the creator of these scripts has attempted to adhere as closely as possible to the conventions of Hokkien theatre. In doing so, a unique fragment of the Sinophone lost elsewhere has been preserved in the Philippines.

References

Bai, Yonghua 白勇華, 'Zhanzhuan Dongnanya: Gaojiaxi haiwai bainian 輾轉東南亞：高甲戲海外百年 (1840–1940)', *Fujian luntan, renwen shehui kexue ban*, 8 (2011): 64–68.

Chan Yap, Gloria, *Hokkien Chinese borrowings in Tagalog* (Pacific Linguistics, Series B, 71, Canberra: Australian National University, 1980).

Chen, Zhiping 陳支平 and Xu Hong 徐泓, *Minnan wenhua baike quanshu* 閩南文化百科全書 (Fuzhou: Fujian renmin chubanshe, 2009).

Chia, Caroline, 'Gezai xi in Singapore: Oral transmission, improvisation and dependence on "Fixed Texts"', *CHINOPERL*, 37/1 (2018): 1–41.

Chia, Caroline, *Hokkien theatre across the seas* (Singapore: Springer Singapore, 2019).

"Fei hua xiqu de xingqi 菲華戲曲的興起 (–1945)", in Zhou Ning (ed.), *Dongnanya huayu xijushi* (xiace) 東南亞華語戲劇史（下冊）, vol. 2 (Xiamen: Xiamen daxue chubanshe, 2007), 802–71.

"Fujian sheng difangzhi bianji weiyuanhui" (ed.), *Fujian sheng zhi, xiqu zhi* 福建省志 戲曲志 (Beijing: Fangzhi chubanshe, 2000).

Gonzales, Wilkinson Daniel Ong Wong Gonzales, 'Philippine Hybrid Hokkien as a Postcolonial Mixed Language: Evidence from Nominal Derivational Affixation Mixing' (MA thesis, National University of Singapore, 2018).

Goody, Jack, *Literacy in Traditional Societies* (Cambridge: Cambridge University Press, 1975).

Hung, Wei-jen 洪惟仁, 'Shiliu shiqi shiji zhijian Lüsong de Zhangzhou fangyan 十六、七世紀之間呂宋的漳州方言', *Lishi dili* 歷史地理, 30 (2014): 215–38.

Klöter, Henning, *The Language of the Sangleys: A Chinese vernacular in missionary sources of the seventeenth Century* (Leiden; Boston: Brill, 2011).

Lin, Huadong 林華東, *Quanzhou fangyan yanjiu* 泉州方言研究 (Xiamen: Xiamen daxue chuban she, 2008).

Ministry of Education, Taiwan Minnanyu Changyongci cidian. https://twblg.dict.edu.tw/holodict_new/default.jsp.

Ng, Kim Chew, 'Sinophone/Chinese: "The South where Language is lost" and Reinvented, translated by Brian Bernards', in Shu-mei Shih, Chien-hsin Tsai and Brian Bernards (eds.) *Sinophone Studies: A Critical Reader* (New York: Columbia University Press, 2013), 74–92.

Ng, Percy, 'Kao Ka, the Chinese Opera in the Philippines: A preliminary socio-historical study' (MA thesis, University of the Philippines Diliman, 2016).

Pan, Rongyang 潘榮陽, 'Gaojiaxi yu Min-Tai shehui bianqian 高甲戲與閩台社會變遷', *Dongnan Xueshu* 東南學術, 3 (2013): 219–25.

Ruizendaal, Robin, *Marionette theatre in Quanzhou* (Boston: Brill, 2006).

Shih, Jui-lou 施瑞樓, 'Lukang Juying she Nanguanxi zhi yanchu ji "Yichun liu shan" shenfen zhi tan tao 鹿港聚英社南管戲之演出及〈益春留傘〉身段之探討 (MA thesis, National Taiwan Normal University, 2012).

Shih Shu-mei, 'Introduction: What Is Sinophone Studies?', in Shih Shu-mei, Chien-Hsin Tsai and Brian Bernards (eds.), *Sinophone Studies: A Critical Reader* (New York: Columbia University Press, 2013), 1–16.

Stenberg, Josh, 'Xiqu in the Philippines: From Church Suppressions to MegaMall Shows', *Journal of Chinese Overseas*, 16 (2020): 58–89.

Tan, Sooi Beng and Nancy Yunhwa Rao, 'Introduction – emergent Sino-soundscapes: musical pasts, transnationalism and multiple identities', *Ethnomusicology Forum*, 25/1 (2016): 4–13.

Wickberg, Edgar, 'Hokkien-Philippines familial transnationalism, 1949–1975', in Maria N. Ng and Philip Holden (eds.), *Reading Chinese transnationalisms* (Hong Kong: Hong Kong University Press, 2006).

Wu, Shaojing 吳少靜, 'Minnan yinyue zai Dongnanya de chuanbo 閩南音樂在東南亞的傳播', *Overseas Chinese Journal of Bagui*, 4 (2006): 24–28.

Yu, Anson, 'Fading glory: Chinese opera in the Philippines', *Tulay Fortnightly* (17 April 2007): 8–11.

Yu, Anson, 'Kaoka: A dying art', *Tulay Fortnightly* (October 22 to November 2013): 8–10.

Zhou, Changji 周長楫 (eds.), *Minnan fangyan da cidian* 閩南方言大詞典 (Fuzhou: Fujian renmin chubanshe, 2006).

Zhuang, Changjiang 莊長江, *Quannan Xi Shi Gou Chen* 泉南戲史鉤沉 (Taipei: Guojia chubanshe, 2008).

Zhuang, Changjiang 莊長江, *Quanzhou xiban* 泉州戲班 (Fuzhou: Fujian renmin chuban she, 2006).

Zhuang, Guotu 莊國土, *Feilübin huaren tongshi* 菲律賓華人通史 (Xiamen: Xiamen daxue chubanshe, 2012).

"Do You Love China or Not?": Late-Colonial Textbooks to Learn Mandarin through Malay

Tom Hoogervorst

1   Introduction[1]

On 2 October 1913, roughly two years after the Republic of China was founded, a long letter landed on the desk of the Netherlands Indies governor-general Idenburg. In it, a senior advisor to the colonial government on Chinese affairs, William J. Oudendijk, provided detailed recommendations to improve the education of its Chinese subjects. Government-facilitated opportunities had been disappointingly inadequate, he contended, so that many families opted to send their children to schools oriented to China rather than the Netherlands. This situation was inexcusable if the Chinese were to become full participants of the Dutch Empire. Having studied in Beijing, the seasoned diplomat was furthermore unimpressed with the level of Mandarin taught in the archipelago:

> Not only is Mandarin currently in fashion as a working language, but, amidst the large variety of dialects spoken by the resident Chinese, the large difficulty of choosing one dialect is resolved if Mandarin is adopted. Let me immediately add that in most schools a very peculiar language passes for Mandarin, and that a great many teachers simply just speak their own provincial dialect, sometimes with a little touch of Mandarin.[2]

What did this peculiar type of would-be Mandarin look like, or, rather, sound like? What sources are available to reconstruct its early history in

---

1   Throughout this chapter, I will cite the Chinese data in their original transcription. To identify their contemporary pronunciations, I use *Pīnyīn* for Mandarin and *Pe̍h-ōe-jī* for Hokkien.

2   […] het Mandarijnsch is thans als voertaal niet alleen in de mode, doch bij de groote verscheidenheid der door de hier gevestigde Chineezen gesproken dialecten wordt de groote moeilijkheid der keuze van een dialect doorgehakt zoo men Mandarijnsch neemt. Laat ik hier terstond bijvoegen, dat in de meeste scholen een heel raar taaltje voor Mandarijnsch moet doorgaan, en dat zeer vele onderwijzers maar heel eenvoudig hun eigen provinciaal dialect spreken, soms met een Mandarijnsch tintje er aan (van der Wal 1963: 263).

Indonesia? What can the story of Netherlands Indies Mandarin tell us about Chinese-Indonesian history more broadly?

Roughly speaking, the history of Mandarin in Indonesia exhibits two distinct periods. The first started at the turn of the twentieth century, when efforts to modernize Chinese education – both in China and among the Chinese-descended communities in Southeast Asia – led to the widespread adoption of Mandarin as the language of instruction at the expense of pre-existing varieties such as Hakka, Cantonese, Teochew, and especially the Zhangzhou variety of Hokkien. In the Netherlands Indies – as Indonesia was known under Dutch colonialism – the Chinese Meeting Hall or *Tiong Hoa Hwe Koan* 中華會館 played a crucial role in the establishment of modern Mandarin-medium schools (Suryadinata 1972; Sai 2016; Kwartanada 2018). Mandarin continued to be taught under the Japanese occupation (1942–45) and in Sukarno-era Indonesia (1945–66). From 1966 to 1998, Chinese education and linguistic expressions were banned as part of the assimilationist regime of Indonesia's second president, Soeharto, leading to a period of cultural disruption and destruction (Heryanto 1998; Sai 2006). The second period of Mandarin education in Indonesia, hence, only started after 1998, when the ban on Chinese education was lifted and numerous Chinese-descended Indonesians began to reconnect with their Chinese identity (Sai 2010; Setijadi 2015).

The present chapter deals with the first period – roughly the first half of the twentieth century – which to my knowledge has not yet received much linguistic attention. It investigates the way Chinese-Indonesians in the late-colonial period learned Mandarin, the specific type of Mandarin they were taught, and the material they used. By comparing several textbooks that relied on Malay/Indonesian as the language of instruction, I will examine, at once, the linguistic characteristics and the contents of the teaching material. Unlike "Malaysian Mandarin" or "Singaporean Mandarin" (Goh 2017), the Indonesian variety of Mandarin was never broadly adopted even amongst the Chinese themselves. Nevertheless, its linguistic features tell a compelling story of Chinese-Indonesian cultural and political contestations during a pivotal part of their history. The academic appeal of Mandarin in late-colonial Indonesia arguably lies in its heterogeneity, having been adopted more than three decades before language planners in China agreed on a unified pronunciation (cf. Sai 2016: 378–79). Some Mandarin teachers in the Indies used orthographic conventions derived from the Dutch language. Intellectually, however, they were relatively free from European structures of education. The contents of their teaching books, especially the example sentences, make it clear that colonialism was incompatible with the aspirations of this generation of pan-Chinese

chauvinists. As I hope to demonstrate, many of the phrases in Mandarin teaching materials reveal insights into the language practices of their authors and intended readership, but also concrete formulations of their worldview.

But first, the concept "Mandarin" merits some clarification. The original European usage of this word corresponds to the Chinese term *guānhuà* 官話 'language of officials' as used during the Ming and Qing dynasties. Especially in the Qing Dynasty, *guānhuà* was a non-standardized prestige variety (koiné) that had emerged from contact between several mutually intelligible dialects. It tolerated a degree of regional variety from *běiyīn* 北音 'northern pronunciations' and *nányīn* 南音 'southern pronunciations', spoken respectively in the Beijing and Nanjing area (Coblin 2007: 23; Simmons 2017: 72). Even with Beijing at the centre of political power, most literati preferred the prestigious southern variety of Nanjing until the early twentieth century (Coblin 2000: 267–68; Simmons 2017: 67; Kuiper 2017: 88 fn. 186). Influenced by Japanese and Western notions of state formation and accompanied by ideas of a Correct Pronunciation (正音 *zhèngyīn*), the concept of a National Language (國語 *guóyǔ*) gradually gained ground among Chinese intellectuals around the turn of the twentieth century (Coblin 2007: 43). In its final years, the Qing government proposed a *guóyǔ* based on the Beijing variety, enriched with some southern Mandarin features (Simmons 2017: 73). A more mixed interdialectal model known as Blue-Green Mandarin (藍青官話 *lánqīng guānhuà*)[3] was promoted by the ROC from 1912, yet this project failed due to the absence of native speakers (Li 2004: 102; Simmons 2017). In the 1920s, support grew for a New National Pronunciation (新國音 *xīn guóyīn*), for which the Nanjing-influenced educated stratum of Beijing Mandarin served as a concrete dialect base. This variety was officially promulgated in 1932 (Li 2004: 103; Coblin 2007: 24; Simmons 2017: 79–82). Concomitant to this development was the popularity of *báihuà* 白話 – the vernacular idiom, as opposed to Classical or Literary Chinese – as a modern written language accessible to the masses (Weng 2018).

The impact of these developments on the history of Mandarin in Indonesia and elsewhere in Southeast Asia remains poorly studied. Many Chinese-Indonesians were more literate in Malay, the archipelago's lingua franca, than in any Sinitic variety. In addition, many could speak and write regional Indonesian languages, such as Javanese and Sundanese. From the early twentieth century, however, the so-called Sino-Malay press regularly called attention to *guóyǔ* and *báihuà* – spelled as *kuo-yü* and *pai hwa* – and the perceived importance for Indies Chinese to learn them. The term *zhèngyīn* – in

---

3   So named because it contained regional and non-standard features and was thus "neither purely blue nor purely green" (Simmons 2017: 63–64, fn. 1).

the archipelago typically pronounced as *tjeng im* (literary Hokkien: *chèng-im*) or *tjia im* (colloquial Hokkien: *chià^n-im*) – was used synonymously with Mandarin. Yet unlike mainland China, Indonesia lacked native-speaking reference points for this language. Job Moerman, an elementary teacher at Chinese schools, observed in 1929 that Mandarin was only used for reading (Moerman 1929: 43). The Chinese-Indonesian journalist Nio Joe Lan 梁友蘭, writing in 1932, noticed that students who had learned Mandarin at school rarely used it at home, where other Sinitic varieties and/or Malay prevailed (Nio 1932: 1092). This situation is reminiscent to what has been observed among Singaporean Chinese in the 1990s – "Mandarin no longer has any relationship to their lived reality. At most it simply triggers a sentimental connection ... the study of Mandarin is actually the study of a foreign language" (Wang 1993 quoted by Ng 2013: 84) – although for many people "no longer has" could be replaced by "never had". Only a very small number of recent migrants from Hunan and northern China reportedly spoke Mandarin natively, and many of them predictably worked as teachers at private Chinese schools (Moerman 1929: 27). Nevertheless, growing numbers of Chinese-Indonesians sought to further their education in China from 1906 (Suryadinata 1972: 62–63; Sai 2006: 151; Hoogervorst 2021: 45), where they must have improved their Mandarin proficiency. On a local level, too, opportunities abounded to study the language (Fig. 7.1). The specific type of Mandarin they learned, hence, merits academic attention.

2 Learning Chinese in Indonesia

Mandarin is Indonesia's only Sinitic language for which courses and textbooks were publicly advertised. In late-colonial times, they were often promoted in the verbiage of pan-Chinese solidarity, typically in Malay (Hoogervorst 2021: 42–50). If one truly supported Chinese nationalism and wished to transcend provincial loyalties, so the argument went, the adoption of a shared national language was the only viable option. Such sentiments could of course easily be exploited for commercial gain. The following advertisement, for instance, inexorably connects a person's love for the Chinese nation to their willingness to purchase a long list of textbooks:

> Do you ...?
> Do you like the Chinese nation?
> Do you want to know about the Chinese civilization?
> Do you want to become a friend of the Chinese nation?

Hsing Chung Hui Kuo Yu Pan di Soerabaja.

Cursus bahasa Tionghoa (Kuoyu) dari Hsing Chung Hui, Soerabaja, jang diboeka pada 1 Januari 1937
dan sampe sekarang masi dapet penoe perhatian.
Foto: Studio „Han" Soerabaja.

FIGURE 7.1   Photo of a Mandarin course in Surabaya
             FROM *SIN PO WEKELIJKSCHE EDITIE* 841 (1939), 5. THE COURSE WAS
             PROVIDED BY THE REVIVE CHINA SOCIETY OR *HSING CHUNG HUI* 興中會.
             NO KNOWN COPYRIGHT HOLDERS

> Do you plan to go to China?
> Do you understand the Mandarin Chinese language?
> Do you enjoy learning the Chinese script?
> Do you want to learn the Mandarin Chinese language?
> Do you want to interact with the civilized Chinese nation?[4]

Other institutions provided opportunities to learn Mandarin for free. Another advertisement, published in 1924 in the Sino-Malay newspaper *Tjin Po* 貞報, invokes the perennial bugbear of cultural attrition. As we can read below, the Sino-Malay newspaper *Keng Po* 競報 offered a free course according to the system created by a certain Tjiam Dji Ko, of whom I have not been able to find any further information:

---

4 Apatah …? | Apatah Toean soeka sama bangsa Tionghoa? | Apatah Toean ingin kenal kasopanan Tionghoa? | Apatah Toean maoe djadi sobatnja bangsa Tionghoa? | Apatah Toean aken pergi ka Tiongkok? | Apatah Toean mengarti bahasa Tionghoa Tjeng-Im? | Apatah Toean soeka berladjar hoeroef Tionghoa? | Apatah Toean maoe berladjar bahasa Tionghoa Tjeng-Im? | Apatah Toean ingin bergaoel dengan bangsa T.H. jang sopan? (Tip 1923).

> Learn the Chinese script for free. Seeing as the Chinese schools have failed to become places to prevent the loss of the Chinese language and script among the Indies-born Chinese, starting from 8 April, *Keng Po* in Batavia will teach five Chinese characters in the Mandarin dialect in each issue, totalling 1000 characters according to Tjiam Dji Ko's system. [...] Support the goal to make Chinese the language of social interaction in the Indies.[5]

The best-rated Chinese schools in the Netherlands Indies managed to offer their entire curriculum in Mandarin, without relying on other languages (Sai 2016: 384). Exclusively Chinese teaching materials were not romanized, although some had phonetic markings known as *Zhùyīn Fúhào* 注音符號. For more detail on the phonology of Indonesia's early Mandarin, therefore, we must direct our attention to commercially published phrasebooks and dictionaries with explanations in Malay, which the majority of learners would have required. Below, I will briefly highlight the background and contents of this chapter's primary sources.

The first commercially available book known to me from which Indies-born Chinese could hone their Mandarin skills was a 140-page wordlist of Malay, Dutch, Japanese, Mandarin, Hokkien, Hakka, and Cantonese titled *East Asia*, first published in 1910 by John F. Knoetsen.[6] Not much is known about the book's compiler, who must have been born around 1873 and worked as a telegraphist and non-military clerk (*burgerschrijver*) in Palembang, a city in Sumatra. Some of the photographs Knoetsen took of Palembang life at the beginning of the twentieth century are kept at the Leiden University Library.[7] Netherlands Indies newspapers report that he took a one-year furlough to China in October 1905, where he must have nurtured his Mandarin proficiency. In 1906, the self-taught polyglot briefly taught English at the Tiong Hoa Hwe Koan school in Batavia (Nio 1940: 93–94), which made him a contemporary at that institute of the famous Singaporean educational reformist Lim Boon Keng 林文慶. Next to *East Asia*, Knoetsen published an English-Malay-Javanese-Sundanese phrasebook (1911), a manual to read and

---

5 Bladjar hoeroef Tionghoa dengen prodeo. Meliat Haktong telah loepoet mendjadi tempat aken tjega terhilangnja hoeroef dan bahasa Tionghoa antara pranakan-pranakan di Hindia, maka moelaï dari tanggal 8 april KENG PO di Batavia setiap terbitnja ada mengadjar lima hoeroef Tionghoa dalem dialect Tjeng Im sampe 1000 hoeroef systeem Tjiam Dji Ko. [...] Toendjanglah ini maksoed boeat bikin pergaoelan di ini Hindia mendjadi Tionghoa.

6 I have not been able to find the first edition. Copies of the third edition, published in 1912, can be found at the universities of Leiden and Cornell.

7 KITLV 4784, KITLV 180707, KITLV 181505.

write Chinese characters (1931),[8] a Dutch-Japanese-Malay-English phrasebook for tourists (1934), and a book containing templates for legal correspondence (1935). These were all published by the Chinese-owned printing house Tjiong Koen Bie 鍾昆美 in Batavia, later known under the name *De "Pertoendjoengan"*. Another source of income for Knoetsen, both before and after his retirement in 1919, was the writing of endorsements for pharmaceutical companies; his name can be found in various advertisement for pills against rheumatic pains, haemorrhoids, and other ailments. No amount of medicine could save him, however, when he got hit by a car in 1938 and died at the age of 65, as several Netherlands Indies newspapers reported.

Unlike the other books discussed in this chapter, *East Asia* contains only words and no example sentences. Its publication by Tjiong Koen Bie and the laudatory forewords of Chinese dignitaries – the Qing representative Ong Hong Siang 王鳳翔 and the Chinese newspaper editor Pe Pin Chuw 白萍洲 – suggest that Knoetsen was well-established in Chinese circles. As becomes clear from the introduction of the book's second edition (1912), his targeted readership consisted of Chinese and Europeans alike:

> Since I have started to publish my little "Handbook", I have noticed that not only local-born Chinese, but also various Europeans are aficionados of those "Handbooks", even though they are chiefly written in Malay; and I have also realized that it is very useful for Dutch traders, travellers, missionaries, and police servants in the Netherlands Indies to learn to speak Japanese or one of the Chinese languages.
>
> That is why I have decided to completely revise the second edition, in such a way that a full-blooded Dutchman, who does not know a single word of Malay, would be able to find with great ease a Dutch word translated into Japanese or Chinese.[9]

---

8   I have not been able to find any copies of this publication.

9   Sedert dat ik begonnen ben met myn "Handboekje" uit te geven, heb ik bemerkt dat niet alleen hiergeborene Chineezen maar ook verscheidene Europeanen liefhebbers zyn van die "Handboekjes", niettegenstaande de hoofdtaal daarvan Maleisch is; en ook heb ik ingezien, dat het leeren spreken van Japansch of een der Chineesche talen daarin aanwezig zeer nuttig is voor Hollandsche handelaren, reizigers, zendelingen en politie-ambtenaren in Ned: Indië. | Daarom heb ik besloten de tweede uitgave geheel te wyzigen, en wel zoodanig, dat een volbloed Nederlander, die nog geen enkel Maleisch woord kent, zeer gemakkelyk een Hollandsch woord in het Japansch of Chinees vertaald, zou kunnen vinden. (Knoetsen 1912: ii).

Also in 1912, a Chinese-English-Malay phrasebook titled *The Chinese, English and Malay Classified Conversations* was published by Chun Lim &Co 振林公司 in Batavia.[10] I have not been able to find anything about its compilers: K.S. Chu or K.S. Tjoe, S.L. Kim, and B.S. Lim. I suspect the latter was Lim Bok Sioe 林木秀, a close associate of the Tiong Hoa Hwe Koan and the Sino-Malay newspaper *Sin Po* 新報. Judging from some of the example sentences, not all editors were fluent in Malay. Most of the chapters prepare the Indies-born reader for a stay in China. They include conversations about carriages, railroad transportation, ships, custom houses, hotels, and leisure options one was anticipated to enjoy in the "homeland". The book's English introduction made the point that knowledge of Mandarin was essential for the ethnic Chinese residing in Southeast Asia – at the expense of their pre-existing linguistic repertoires – to become part of the Chinese nation:

> In China, the Mandarin language is chiefly used either in the Northern or Southern China; but most of the Chinese, in the Southern island or in the Malay Archipelago, frequently use the Malay language, or some other languages such as Cantonese or Hokianese [sic!], and by continuing these languages so frequently and being so accustomed to them, that they have entirely forgotten one principal language – the language chiefly spoken amongst the Chinese, inhabitants of China and that is the Mandarin language
>
> CHU ET AL. 1912: i

A third important publication appeared in 1920 under the title *Combining Chinese and Malay* or *Hoa Woe Ho Pie* 華巫合璧.[11] This phrasebook was written by the Meixian-born Chun Foo Chun 陳撫辰, who had settled in Batavia in 1906 to engage in educational work among the Indies Chinese. From 1911, he also promoted several Chinese commodities in the Netherlands Indies (CWR 1936: 27). Of the 7 language manuals and 10 Malay novels Chun Foo Chun reportedly authored (ibid.), the majority have gone lost. His 1930 *New Malay-Chinese Dictionary* or *Sin Woe Hoa Tze Tian* 新巫華字典, of which a copy is kept at the National Library of Singapore, is worth mentioning. In addition, Chun Foo Chun wrote short manuals on Chinese medicine (1915) and, as he phrased it, on important matters women should know (1917). Politically, Chun was a formal

---

10   Malay: Boekoe Omong-Omongan Njang Teratoer Darie Bahasa Tjina, Inggris dan Malayoe. The only copy of this publication known to me is kept at the Museum of Chinese-Indonesian Literature (Museum Pustaka Peranakan Tionghoa) in Tangerang, Indonesia.

11   This publication has recently been digitized by Cornell University.

representative of the Chinese Republican government. Commercially, he was associated with J. Lowe Medicine Co. 羅威藥房 in Shanghai and – together with Lim Bok Sioe – with the aforementioned publishing house Chun Lim &Co. Due to Chun's close links with Chinese journalists and educationists in the Indies, advertisements for his *Hoa Woe Ho Pie* can be found in Sino-Malay newspapers and Mandarin teaching books used in Chinese schools (Fig. 7.2).

Most examples in Chun's *Hoa Woe Ho Pie* are specific to the experiences of the Indies Chinese. Its conversations generally take place in Batavia and specifically in the neighbourhood Pasar Baroe 'New Bazaar', which feature in the book as respectively *Pa ta wei* 巴達維 (Bādáwéi) or *Pa tjeng* 巴城 (Bāchéng) and *Sen pa sa* 新巴杀 (Xīnbāshā) 'New Bazaar'. In the introduction, Chun Foo Chun underscores the importance of a book about the Mandarin language:

> I have indeed deliberately written this book for the needs of all you Sirs, who only understand, or know, the shape of Chinese characters, but do not understand what they mean; in this book all of their meanings will become clear; I have divided all that is written in *Hoa Woe Ho Pie* into four separate books.[12]

Of these four volumes, only the first has been preserved.[13] The example sentences of Part One consist of various topics for everyday dialogues. The second volume reportedly focuses on interactions with important people, business talk, and domestic affairs. The third volume delves into Chinese characters with deeper meanings and their use in conversations, whereas the fourth volume deals with written correspondence. While the book was published by Chun Lim &Co in Batavia, it was printed in China:

> This book is of great necessity to obtain for you students in the Netherlands Indies, and for you Sirs who would like to know languages and should also know your own language.
>
> If any Malay letter is misspelled, I offer my sincere apologies, as this book has been printed by typesetters in Shanghai, none of whom can speak Malay.[14]

---

12   Saia sengadja mengarang ini boekoe memangnja boeat keparloean pada L.W. Sian seng, jang tjoema mengarti, atawa kenal, sadja roepanja hoeroep tjina, tetapi tida tahoe menartinja, di dalam ini boekoe bisa menarangkan sekalian maksoednja, soerat jang ada tertoelis di dalam ini Hoa Woe Ho Pie, semoea saja pitjah 4 kepeng boekoe (Chun 1920: ii).

13   The only copy known to me is kept at the Leiden University Library.

14   Ini boekoe ada besar kaperloean boeat di mempoenjain oleh liatwi Hakseng[2] di hindia, balanda [sic!] ini dan boeat liatwi sian seng jang ada hati kapingin tahoe bahasa dan

FIGURE 7.2   A hand-written advertisement of *Hoa Woe Ho Pie*

The fourth book examined in detail here is *Kuo-Yü Zonder Goeroe* 'Mandarin without a Teacher', published in 1934 by a certain Mrs. The (Fig. 7.3).[15] Born in 1916, the author worked as a private Chinese teacher in Surabaya from the mid-1930s. Her *Kuo-Yü Zonder Goeroe*, however, was intended for self-study. It was published by the Surabaya-based accountancy firm New China 新中華, founded in 1926 by her husband The Chung Shen 鄭叢森. In 1937, as several Netherlands Indies newspapers inform us, this Mandarin course was broadcasted across Central and East Java by the Netherlands Indies Radio Broadcasting Corporation (NIROM). A second volume of *Kuo-Yü Zonder Goeroe* appeared in 1940. In 1941, Mrs. The's course appeared in parts in *Eh Thoeng* 兒童 'Child', a journal for Dutch-speaking young Chinese. In 1942, she also published a teaching book on the Chinese alphabet titled *The National Alphabet for Overseas Chinese*.[16]

In post-independence Indonesia, Mrs. The worked with her husband at the Lian Huo 聯合 High School and at his printing house in Surabaya, which published three newspapers: *Chinese Daily News* (in Chinese), *De Vrije Pers* (in Dutch), and *Java Post* (in Indonesian).[17] In 1953, Indonesia-based Dutch newspapers report that she visited western Europe to study women in journalism. The couple's journalistic career faced a major setback in 1956, when their office was vandalized by a mob, leading to widely reported protests by Indonesian journalists. Their newspaper *Java Post*, which still exists today as *Jawa Pos*, proved most durable. Mrs. The had adopted the Indonesian name Mega Endah – and her husband became Suseno Tedjo – yet remained affectionately known as *Tante The* 'Auntie The' (Wangkar 2001: 15). In addition to the pro-ROC stance she evidently adopted sometime between the first and second volume of *Kuo-Yü Zonder Goeroe*, she quickly embraced the Indonesian nation-state and dedicated the rest of her career to the *Jawa Pos*. In the early 1980s, when her husband was in his late 70s and their three children had moved to Europe, she effectively managed this newspaper on her own (Iskan 1982: 28).

---

bahasa sendiri poen perloe djoega. | Kaloe salah swatoe hoeroep melajoe ada tersalah, saia banjak minta maaf sebab ini boekoe di tjitak di Sianghay leter zetter semoa tida tahoe omong melajoe (Chun 1920: iv).

15   Chinese: 鄭夫人, Malay: Njonja The. Her full name was Mrs. The Chung Shen-Tjia 鄭叢 森夫人.

16   Malay: *Boekoe beladjar alphabet Kuo-Yin*, Chinese: 華僑國音字母拼音法.

17   From 1948, *Chinese Daily News* was known as *Hua Chiao Sin Wen* or *Hwa Chiau Hsin Wen* 華僑新聞 'Overseas Chinese News'. At the time it was published, it was said to be Indonesia's only newspaper in Chinese that was politically pro-Indonesian (Iskan 1982: 28). De Vrije Pers ('The Free Press') was eventually renamed to *Indonesian Daily News*.

18392

f 2.60

KUO - YÜ ZONDER GOEROE

SYSTEEM 1000 HOEROEF DENGEN DITERANGKEN DALEM BAHASA BELANDA DAN MELAJOE.

Djilid ka I.

DIKARANG OLEH :

Mevr. THE CHUNG SHEN - TJIA

TJITAKAN KA-AMPAT
SEPTEMBER 1940

DITERBITKEN OLEH :
N. V. ACCOUNTANTSKANTOOR NEW CHINA
SOERABAIA·JAVA.
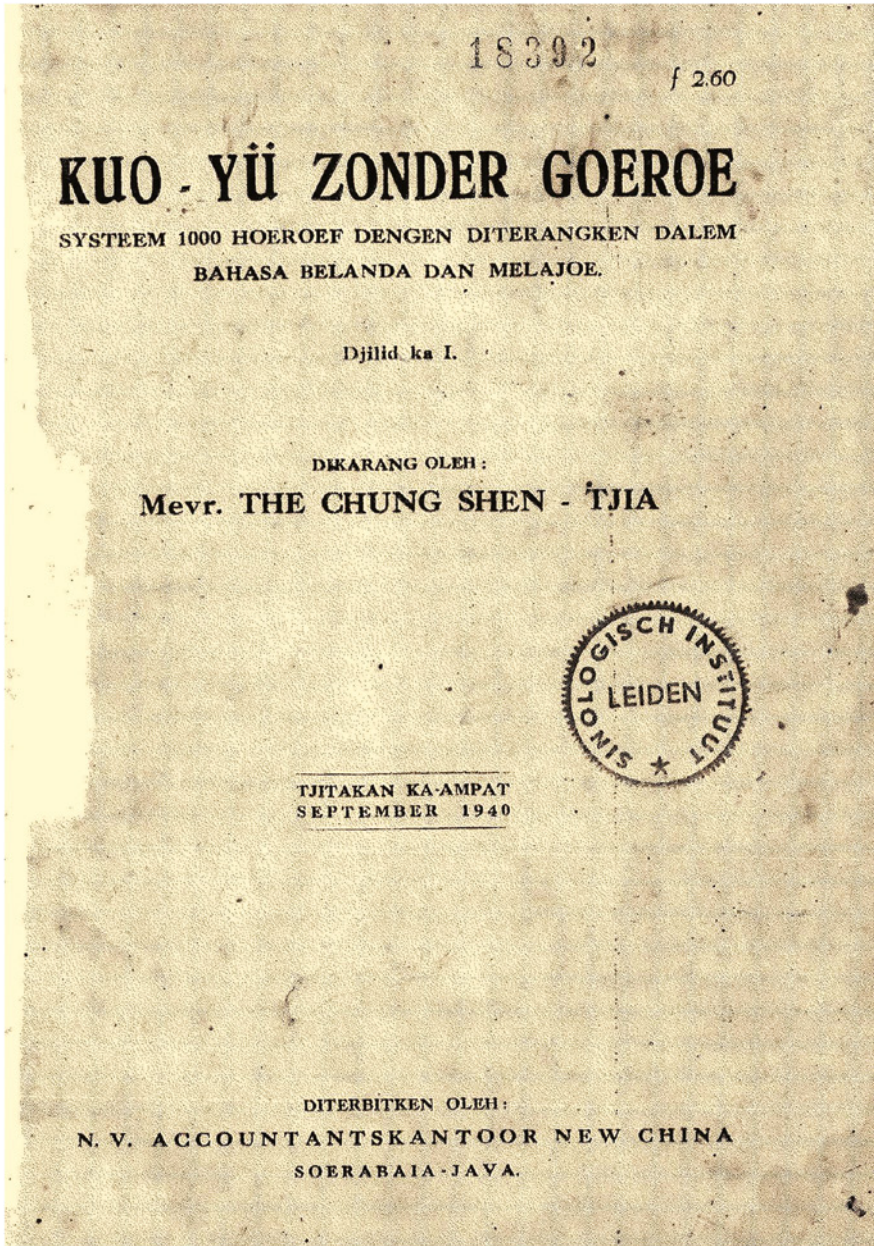
FIGURE 7.3 Cover of Kuo-Yü Zonder Goeroe

In the first part of *Kuo-Yü Zonder Goeroe*, Mrs. The explains that Chinese, like Malay, has less morphology than European languages and is therefore easier to learn. For this book, she developed her own method to learn Mandarin, which required the student to learn the pronunciation of individual characters before having to study their usage in combinations, peruse example sentences, and carry out exercises. To indicate how Chinese characters should be pronounced, Mrs. The used a Dutch-influenced Indies-style romanization alongside Zhùyīn Fúhào transcriptions. In the second part of *Kuo-Yü Zonder Goeroe*, conversely, the entire source text – rather than just the individual characters at the beginning of each chapter – was given in romanized form underneath the Chinese characters. Part Two is remarkably more ideological in nature; it discusses trips to China and the embarrassment of ethnic Chinese who cannot understand "their own language". Several example sentences also mention Toean Cheng or Chěng hsiën shěng 鄭先生, clearly referring to Mrs. The's husband.

The above publications are by no means Indonesia's only educational texts on Mandarin. Next to the aforementioned exclusively Chinese materials, several small Mandarin-Malay dictionaries appeared from the 1910s onwards. One example is the *Chinese-Malay Dictionary* or *Han Woe Tzi Tian* 漢巫辭典, published in 1915 by the Batavia-based printing house Kho Tjeng Bie 高正美. Among its compilers are listed: Eric H. Tshen of the Republican University in Shanghai, Oey Tjiang Hoay 黃昌懷 from Hamburg, and the aforementioned Lim Bok Sioe from Batavia. Chun Foo Chun is also thanked in the introduction. A Mandarin-Malay-Dutch-English dictionary titled *A Classified Dictionary of Chinese-Malay-Dutch-English with Chinese Pronounciation* [sic!][18] was published in 1931 by the National Book Office 國民書局 in Batavia. Its compiler Li Joek Koey 李毓愷 published several more lexicographical works in the 1950s and 1960s. A third dictionary, the *New Chinese-Malay Dictionary*,[19] appeared in 1935 by the hands of the aforementioned Oey Tjiang Hoay, by then a Semarang-based entrepreneur.

In November 1945, Yoe Wan Fei published another Mandarin self-study book titled *Lim's Mandarin without a Teacher*,[20] once again encouraging Indonesia's Chinese community to learn the language and script of their "real" country. Among its targeted readership, as the introduction makes clear, were people

---

18   Chinese: 綜合華巫荷英大辭典: 國語注音; Malay: *Kitab-logat jang mengatoerken dari bahasa-Tionghoa-Melajoe-Olanda-Inggris dengan soeara batjanja dalem bahasa-Tionghoa*; Dutch: *Chineesch-Maleisch-Hollandsch-Engelsch classificeerend-woordenboek met Chineesche uitspraak.*

19   Chinese: 四角號碼華巫新辭典, Malay: *Kitab Logat Baroe Tionghoa Melajoe.*

20   Chinese: 林民國語自修讀本, Malay: *Lim's Boekoe Peladjaran Kuo-Yü Zonder Goeroe.* The only copy known to me is kept at the Leiden University Library.

who had learned some Mandarin at the Chinese elementary school – an institution designated with the curious Hokkien-Mandarin compound *tjoe-têng siao-hsüeh* (*chiūⁿ-tn̂g* 上堂 and *xiǎoxué* 小學) – but did not have a chance to further their education. Published in Jakarta (formerly Batavia) by Lim's Boekhandel, all the book's dialogues take place in China. Very little is known to me about the author, who also occasionally wrote about ancient Chinese history in the aforementioned periodical *Sin Po*.

3 Mandarin, but Which Kind?

From the mid-nineteenth century, Malay newspapers in the Netherlands Indies used a Dutch-influenced Indies-style romanization, rather than the Arabic-derived script in which native Malay speakers wrote their language. While far from consistently applied, these conventions were in broad use, especially in Java's cities. A similar system emerged almost organically for Mandarin when the need to transcribe it first arose. Some of its most salient characteristics are outlined in Table 7.1:

TABLE 7.1   The Indies-style transcription of Mandarin

Romanization	Phonological value	Example
ai, aij, aj, ay	/ai̯/	tsai, tsaij, tsaj, tsay 在 (zài)
ao, au, auw, aw	/au̯/	yao, yau, yauw, yaw 要 (yào)
dj	/d͡ʒ/ ~ /ʐ/	djan 然 (rán)
e, ê, ë, ĕ, ẽ	/ə/	hen, hên, hën, hĕn 很 (hěn)
é, ee	/e/ ~ /ei̯/	tjé, tjee 這 (zhèi)
è, èh	/ɛ/	hsüè, hsüèh 血 (xuè)
ê, i, ŭ	/ɥ/	tzê, tzi, tzŭ 子 (zi)
euw, eoe	/ɤu̯/	tjeuw, tjeoe 走 (zǒu)
ie, ieh, i, î	/i/	tie, tieh, ti, tî 的 (dì)
ni, nj	/ni/ ~ /ɲ/	niauw, njiauw 鳥 (niǎo)
o, ô, òh	/ɔ/	kwo, kwô, kwòh 國 (guó)
oe, oo	/u/ ~ /ʊ/	sioeng, hsioong 兄 (xiōng)
oo, o	/o/	tjoo, tjo 作 (zuò)
oouw, ouw, ou	/ou̯/	thoouw, thouw, thou 頭 (tóu)
sj	/ʃ/ ~ /ʂ/	sjan 山 (shān)
tj	/t͡ɕ/ ~ /t͡ʂ/	tjoeng 中 (zhōng)
ü, uu	/y/	hsü, hsuu 許 (xǔ)

In the above system, the ⟨h⟩ was used to mark aspirated consonants, e.g. *khoeng* 孔 (kǒng), *phei* 陪 (péi), *tha* 他 (tā), and *tjha* 茶 (chá), although this was not always done consistently or correctly.[21] Vowel diacritics conveyed additional phonological detail; the diaeresis ⟨ö⟩ indicated that the underlying vowel formed a distinct phonological unit, e.g. *sjoeï* 誰 (shuí), *tjië* 街 (jiē), and *tjoöe* 稠 (chóu). Other authors preferred the acute accent ⟨ó⟩ for this purpose. Some orthographic idiosyncrasies indeed depended on individual preferences. Chun's *Hoa Woe Ho Pie*, for instance, frequently used ⟨v⟩ to transcribe the sound /f/. The transcription of Mandarin onglides was also inconsistent; *yī* 一 can be found as *i, ie,* or *yi, wù* 惡 as *oe, oeh,* or *woe,* and *yù* 預 as *ü, üh,* or *yü.* The Mandarin retroflex (/ɻ/ ~ /ʐ/) was spelled as ⟨yr⟩ word-initially but ⟨rl⟩ word-finally by Chu et al. (1912), as ⟨jr⟩ or ⟨j⟩ word-initially and ⟨rl⟩ word-finally by Mrs. The (Shen-Tjia 1934), and as ⟨dz⟩ or ⟨dj⟩ word-initially and ⟨l⟩ word-finally by Chun Foo Chun (1920). Fei (1945) transcribed it as ⟨r⟩, but clarified in his preface that it "has to be voiced, with the tongue curled upwards".[22] The first volume of *Kuo-Yü Zonder Goeroe* transcribed Mandarin words as in Table 7.1, yet the second volume adopted a system closer to the Wade-Giles Romanization.[23]

Another source of complication was the inability of most Chinese-Indonesians – in the good company of numerous speakers in China and Southeast Asia – to pronounce retroflex consonants. Dental, palatal, and retroflex consonants are often merged in the Mandarin of Java-born Chinese (Oetomo 1987: 81). This merger seems to have existed in late-colonial times too, judging from the identical way these sounds were transcribed. In early Malay textbooks, the voiceless alveolar sibilant affricate /t͡s/ ⟨z⟩, the voiceless alveolo-palatal sibilant affricate /t͡ɕ/ ⟨j⟩, and the voiceless retroflex sibilant affricate /t͡ʂ/ ⟨zh⟩ were all routinely transcribed as ⟨tj⟩, although the former was also sometimes transcribed as ⟨tz⟩, ⟨ts⟩, or ⟨tjz⟩. Likewise, the aspirated voiceless alveolar affricate /t͡sʰ/ ⟨c⟩, the aspirated voiceless alveolo-palatal sibilant affricate /t͡ɕʰ/ ⟨q⟩, and the aspirated voiceless retroflex sibilant affricate /t͡ʂʰ/ ⟨ch⟩ were all transcribed as ⟨tjh⟩. The Dutch romanization was furthermore unable to accurately reflect the Mandarin sibilants, as was already pointed out by the aforementioned Nio Joe Lan in the 1930s (Nio 1933: 412). The orthographic ⟨s⟩ was used for the voiceless alveolar dental sibilant /s/, the voiceless

---

21   The Philippine Hokkien data in Chia (this volume) are similar in this regard.

22   […] haroes diboenjiken satengah soeara, dengen lida ditekoek ka atas.

23   Dutch Sinologists, conversely, had their own way of transcribing northern Mandarin, in addition to having to know the French, English, and German systems (cf. Kuiper 2017: 100). Unlike Mrs. The's system, the Dutch Sinologists used diacritic markers to indicate tones.

alveolo-palatal sibilant /ɕ/, and the voiceless retroflex sibilant /ʂ/, correspond-
ing respectively to ⟨s⟩, ⟨x⟩, and ⟨sh⟩ in modern Mandarin. However, /ɕ/ was also
occasionally transcribed as ⟨hs⟩, and /ʂ/ as ⟨sch⟩, ⟨sh⟩, or ⟨sz⟩.

The fact that the first volume of *Kuo-Yü Zonder Goeroe* gives Zhùyīn Fúhào
transcriptions alongside Indies-style romanizations affords some further
insights into the way Mandarin was pronounced. Syllables ending in ⟨ㄧㄥ⟩,
which would suggest the pronunciation /iəŋ/, are consistently spelled as ⟨ing⟩
in the Indies-style romanization. This implies that the orthographic schwa
was not supposed to be pronounced in this context. This situation resembles
that of contemporary Taiwan, where the historically common pronunciation
of /iəŋ/ is considered uneducated or non-standard and ⟨iŋ⟩ is systematically
taught at schools (Li 2004: 115). For the syllabic consonant /ʐ̩/, the Zhùyīn
Fúhào transcription implied that the characters 只, 址, 枝, 織, and 知 had to be
pronounced identically as ⟨ㄓ⟩ /ʈʂ̩/, yet in the Indies-style romanization they
were spelled respectively as *tji, tsê̌, tsŭ, tzê̌*, and *tzŭ*. It is difficult to determine
whether this discrepancy reflects the inability of the Indies-style romaniza-
tion to accurately reflect the sounds of Indonesian Mandarin, or of the Zhùyīn
Fúhào system to accurately reflect its minute phonological differences.

With diacritics already serving other purposes, the Indies-style romaniza-
tion was unable to mark tones. Only the first volume of *Kuo-Yü Zonder Goeroe*
contains information about tones. We can see the same four tones as in modern
Mandarin; the fifth 'checked' tone of some earlier standardizations is absent.[24]
As the tones were given for each syllable individually, the book contains no
indication of a neutral tone, as attested in modern Mandarin from mainland
China (but not Taiwan). Even the particles *le* 了 (le) and *tzŭ* 子 (zi) were listed
as carrying the third tone. The other three textbooks consistently spelled the
particles *ti* 的 (di ~ de), *liao* 了 (liao ~ le), and *ni* 呢 (ni ~ ne) in their "full" rather
than unstressed "schwa" forms. In Chun's *Hoa Woe Ho Pie*, we can occasionally
observe an orthographic ⟨h⟩ marking the falling tone, e.g. *mhai* 賣 (mài) 'to
sell' vs. *mei* 買 (mǎi) 'to buy' and *nah lie* 那裡 (nàlǐ) 'there' vs. *na lie* 哪裡 (nǎlǐ)
'where', but this usage is not consistent. In fact, the author contended that
proper pronunciation could only be learned through conversation, although
he remained open to feedback from his readers on this matter:

> As has been said previously, this book does not use tone-indicating let-
> ters as this would increase the burden on its user – this is done in most
> books, adding to the confusion of the learner – because to teach pro-
> nunciation requires continuous familiarization with speech in order to

---

24   This tone is known in Chinese linguistics as the "entering tone" (入聲 *rùshēng*).

quickly soften the tongue, alongside attention to easy, difficult, long, and short conversations, so that the listener will be able to understand them quickly; this is the most important thing, but in this book I will try to do a bit of both. You students may consider for yourselves what works best, and please let me know your thoughts, so that the third print can be adjusted to what most people want.[25]

Another complication was the aforementioned coexistence of a northern and a southern variety of Mandarin, known under different names. In Knoetsen's *East Asia*, they were introduced as Kiangsoe 江蘇 and Pitjili 北直隸 Mandarin, i.e. the Jiangsu (Nanjing area) and North Zhili (Beijing area) varieties. Knoetsen's summary of their sound correspondences is given in Table 7.2 (he provides no characters or examples). Throughout his dictionary, both pronunciations are juxtaposed if different.

TABLE 7.2   Kiangsoe 江蘇 and Pitjili 北直隸 correspondences

Kiangsoe	Pitjili
eo	ou
kuu	tjuu
khuu	tjhuu
ki	tji
khi	tjhi
r	dj
o	oe
oei	eeï
ioo	oéé

DATA FROM KNOETSEN (1912: IV)

---

25   Seperti soedah dibilang tadi ini boekoe tida memake hoeroep pemantes swara memang djoega djadi lebeh membratkan kapada jang memake, begitoelah boekoe² jang kebanjakakkan hingga menambah bingoengnja jang adjar, sebab adjar omong perloe sadja saban waktoe di biasaken beromong² djadi lidah lekas lembek dan perhatikan djoega omongan jang enteng atawa brat pandjang atawa pendek, djadi bisa lekas mengarti orang jang mendengar, itoe jang paling perloe, tetapi di boekoe ini saia tjoba djoega tjampoer sedikit. Liatwi siangseng boleh timbang sendiri jang mana dikira baek harap di kabari sedikit pada saia, soepaia tjapan jang katiga bisa tjoetjoek menoeroet kemaoean orang banjak (Chun 1920: 101).

In the second volume of *Kuo-Yü Zonder Goeroe*, it is clarified that the Mandarin of northern China – *Pe-fang kuo-yü* 北方國語 – will be used. As the author explains, this is the pronunciation of Beiping 北平.[26] Nevertheless, she also calls attention to a number of archaic pronunciations that prevail among the Nanyang Chinese (Table 7.3). These, she explains, correspond with previously learned equivalents from the Mandarin of southern China:

TABLE 7.3   North Chinese and Nanyang correspondences according to Mrs. The

Southern Mandarin	Northern Mandarin	Character	Meaning
ngo	wo	我 (wǒ)	I
pe	pai	白 (bái)	white
hsio	hsuéh	學 (xué)	to learn
jin	rên	人 (rén)	person

DATA FROM (SHEN-TJIA 1940: II)

In the first volume of *Kuo-Yü Zonder Goeroe*, too, we already find a predominance of northern Mandarin features. The rhotacization of syllable finals, for example, was (and is) more common in the Mandarin of Beijing than of Nanjing.[27] We indeed encounter the forms *na'rl* 哪兒 (nǎr) 'where', *tjê'rl* 這兒 (zhèr) 'here', and *yi tia'rl* 一點兒 (yīdiǎr) 'a little'. In the *Classified Conversations*, too, these occur as *na erl*, *tjee erl*, and *i tien erl*.

The substitution of southern for northern Mandarin, however, was a gradual process. Yoe Wan Fei advised as late as 1945 that Mandarin ⟨ou⟩ /oṳ/ had to be pronounced as in Sundanese: ⟨eu⟩ /ɤ/. This older pronunciation resembles that of southern Mandarin (Coblin 2000: 301, 327) and can also be seen in the other textbooks, e.g. *theuw* 頭 (tóu) 'head', *tjeuw* 走 (zǒu) 'to walk', and *tjheoe* 醜 (chǒu) 'ugly'. In other attested vocabulary, the ⟨ou⟩ of standard Mandarin corresponds to the high back rounded vowel /u/ in early Indonesian textbooks, e.g. *pêng yoe* 朋友 (péngyǒu) 'friend', *toe* 都 (dōu) 'all', and *yoe* 有 (yǒu) 'have'. Another feature seen in the textbooks is the innovation *wɔ > ɔ, which took place in southern but not in northern Mandarin dialects (Coblin 2000: 317–18), e.g. *sho* or *sòh* 說 (shuō) 'to say', *tjoh* 錯 (cuò) 'wrong', and *to* 多 (duō) 'many'. We also frequently encounter a historically common ⟨o⟩ where standard Mandarin

26   Beiping was the name of Beijing from 1928 to 1949.
27   This phonological process is known in Chinese linguistics as *érhuà* 兒化.

now has ⟨e⟩, e.g. *koh* 割 (gē) 'to cut', *kho ie* 可以 (kěyǐ) 'can', and *na ie ko* 哪一個 (nǎyīgè) 'which one'. The latter two features remain common in the informal Mandarin of southern China and Taiwan. Some additional "pre-standardized" pronunciations are listed in Table 7.4:

TABLE 7.4   Archaic pronunciations in Indonesian Mandarin

Textbooks	Modern	Character	Meaning
hê	hēi	黑	black
i	tā	她ᵃ	she
ke ~ kie	gěi	給	to give
loe	liù	六	six
mo	me	麼	PART
ngay	ǎi	矮	short
ngô	è	餓	hungry
o	ā	啊	PART
pê	bǎi	百	hundred
po	běi	北	north
sê mo	shéme	什麼	what
sjoeï ~ sooi ~ swe	shéi ~ shúi	誰	who
tho	tā	牠	it
tjia fei	kāfēi	咖啡	coffee
tjio	jiǎo	角	10-cent coin
tjoey ~ tjhu ~ ki	qù	去	to go

a   I have not come across the "expected" character 伊 (yī) in the textbooks.

The textbooks also contain some currently non-standard (yet still widespread) vocabulary, e.g. *li pai* 禮拜 (lǐbài) 'week', *tjhai i* 差役 (chāiyì) 'policeman', *tji tjê* 汽車 (qìchē) 'steam tram', and *woe lai* 巫來 (wūlái) 'Malay'.[28] The word for 'don't' is *poe-yau* 不要 (bùyào) in Knoetsen (1912) and Mrs. The (Shen-Tjia 1934), but *moyauw* 莫要 (mòyào) in Chun (1920) and *piêh* 別 (bié) in Chu et al. (1912). *East-Asia* exhibits the words *hang-sji* 行市 (hángshi) 'market', *moe-ju-kièn* 沐浴間 (mùyùjiān) 'bathroom', and *nioe-nai-joe* 牛奶油 (niúnǎiyóu) 'butter', yet *Kuo-Yü Zonder Goeroe* – which came out roughly two decades later – gives

28   The present-day equivalents are *xīngqí* 星期, *jǐngchá* 星期, *zhēngqì diànchē* 蒸汽電車, and *Mǎlái* 馬來, while *qìchē* 汽車 currently refers to a 'car'. Evan Morgan's revised list of Chinese neologisms (1932) exhibits yet another word for 'policeman': *xúntǔ* 巡土.

shang-tjang 商場 (shāngchǎng), *yu-fang* 浴房 (yùfáng), and *huang-yu* 黃油 (huángyóu) for the same concepts.[29] Similarly, *Hoa Woe Ho Pie* displays *tjwo twa tjeg* 腳踏車 (jiǎotàchē) for 'bicycle', whereas *Kuo-Yü Zonder Goeroe* has *tjio-tjhê* 脚車 (jiǎochē). While I have not come across any obvious non-standard characters, the material contains some "variant" characters (allographs), such as 箇 (*ko*) for the numeral classifier 個 (*gè*), 輭 (*djoeaän*) for 軟 (*ruǎn*) 'soft', 裏 (*lie*) for 裡 (*lí*) 'inside', 那末 (*na mo*) for 那麼 (*nàme*) 'so', 甚麼 (*shen mo*) for 什麼 (*shéme*) 'what', and 湮 (*zei*) for 濕 (*shī*) 'wet'. All the textbooks use traditional characters, yet they feature a small number of shorthand variants, e.g. 华 (*hwa*) for 華 (*huá*) 'Chinese', 脚 (*tjio*) for 腳 (*jiǎo*) 'foot', 巴杀 (*pa sa*) for 巴剎 (*bāshā*) 'market', and 雅致 (*ya tzi*) for 雅緻 (*yǎzhì*) 'elegant'.

4   Sino-Malay: A Hybrid Target Language

The textbooks afford similar linguistic observations on the history of Malay. Earlier scholarship has underlined the importance of "the hybridized Malay-Hokkien vernacular" to the study of Mandarin in Indonesia (Sai 2016: 375; Hoogervorst 2021: 42–50). All the textbooks examined here are indeed written in a type of colloquial Malay saturated with Hokkien loanwords, which "behaved" similarly to Baba Malay (Aye, this volume) on a typological level. Yet because both the Malay and the Hokkien dialects differed from those in Malaya, the vernacular that is analysed below should be regarded as a distinct variety. Before delving into its Hokkien elements, it is important to first investigate the Malay itself as used by the Chinese in Java's late-colonial cities. As the primary lingua franca of maritime Southeast Asia, Malay was learned by Chinese visitors and settlers from at least early-modern times. The Netherlands Indies government likewise adopted Malay for administrative purposes, but used a standardization that was quite remote from what was actually spoken by the colony's urban middle-classes. The Mandarin textbooks investigated here all exhibit the Malay used in Java. Even Knoetsen, who must have learned his Malay in Sumatra, used words specific to Java. Table 7.5 below lists some common lexical items found in the Sino-Malay textbooks, along with their equivalents in the "standard" Malay promoted by the colonial government. The grammar of Java's urban Malay has been examined by previous scholars (Oetomo 1991; Mahdi 2016).

---

29   Cf. *shìchǎng* 市場 'market' and *nǎiyóu* 奶油 'butter' in Morgan (1932).

TABLE 7.5   Common Java Malay words in the textbooks

Textbooks	Standard Malay	Meaning
begimana, pigimana	bagaimana	how
belon tau	beloem pernah	never before
boto	tjantik	pretty
kaler	oetara	north
kasoesoe	terboeroe-boeroe	in a hurry
katjek	berbeda	different
kidoel	selatan	south
koedoe	haroes	must
koelon	barat	west
lantaran	sebab	because
lantas	segera	immediately
mama-tjang	nenek	grandmother
ngambang	terapoeng	to float
njang	jang	REL
papa-tjang	kakek	grandfather
perdio	tjoema-tjoema	free of charge
saban	setiap	each
wetan	timoer	east

While many Chinese-Indonesian families had lost fluency in Zhangzhou Hokkien or never spoke it to begin with, their colloquial Malay was permeated with "heritage words" from that language. These words were almost exclusively written in Indies-style romanization. For reasons of recognizability, however, I also give the corresponding Chinese characters and their modern transcriptions throughout this section. In addition to Hokkien loanwords that were widespread even among non-Chinese speakers of Malay – such as *goea* 我 (góa) 'I', *loe* 汝 (lú) 'you', *taotjang* 頭鬃 (thâu-chang) 'queue, pigtail', *Tiongkok* 中國 (tiong-kok) 'China', and *Tionghoa* 中華 (tiong-hoa) 'Chinese' – the textbooks commonly employ such terms as *hakseng* 學生 (hák-seng) 'student', *haktong* 學堂 (hák-tōng) 'school', *liatwi* 列位 (liát-ūi) 'all of you', and *sedji* 細膩 (sè-jī) 'cautious, polite'.

Even more importantly, Hokkien vocabulary was systematically relied upon to translate Mandarin words for which no Java Malay equivalents could

be found. The *Classified Conversations*, for example, translates *miên pao* 麵包 (miànbāo) 'steamed bun', *sing tjen* 生辰 (shēngchén) 'birthday', *soeng pin* 送殯 (sòngbìn) 'funeral procession', and *toeng yang tjê* 東洋車 (dōngyángchē) 'rickshaw' with their Hokkien equivalents *bapauw* 肉包 (bah-pau), *sejit* 生日 (seⁿ-jit), *sangseng* 送喪 (sàng-sng), and *langtjia* 人車 (lâng-chhia). Likewise, the second volume of *Kuo-Yü Zonder Goeroe* translates *tja ts'ai* 什菜 (shícài ~ zácài) 'mixed vegetables' and *tj'a hoe* 茶壺 (cháhú) 'teapot' as *tjap-tjai* 雜菜 (cha̍p-chhài) and *theeko* 茶鈷 (tê-kó·), whereas *Hoa Woe Ho Pie* translates *ta tjhoeen* 打拳 (dǎ quán) 'boxing' and *toe soe* 讀書 (dú shū) 'to study' as *koentaü* 拳頭 (kûn-thâu) and *taktje* 讀冊 (tha̍k-chheh). For the expression *koeng shi koeng shi* 恭喜恭喜 (gōngxǐ gōngxǐ) 'congratulations!', he clarifies that the Hokkiens pronounce it as *kiong hi kiong hi* 恭喜恭喜 (kiong-hí kiong-hí). The Mandarin term *joen toeng* 運動 (yùndòng) 'to exercise' is left as such in Malay (Chun 1920: 59), seemingly because it was well known in Chinese-Indonesian circles. Table 7.6 lists some additional Hokkien-derived translations of Mandarin words attested in the textbooks:

TABLE 7.6 Common Hokkien translations of Mandarin words

Mandarin (as in textbooks)	Sino-Malay	Meaning (in textbooks)
ay ya 優呀 (ài ya)	aya 哎呀 (ai-ia)	(exclamation)
ho-shang 和尚 (héshàng)	hweeshio 和尚 (hôe-siōⁿ)	Buddhist monk
hsién-koean 縣官 (xiànguān)	tikoan 知縣 (ti-koān)	district governor
hsi-thién 西天 (xītiān)	saythi 西天 (sai-thiⁿ)	Heaven
khoeng tze 孔子 (kǒngzǐ)	khong hoe tjoe 孔夫子 (khóng-hu-chú)	Confucius
mê 墨 (mò)	bak 墨 (ba̍k)	ink
Nam Jang 南洋 (nányáng)	Lam Yang 南洋 (lâm-yâng)	Southeast Asia
pî 筆 (bǐ)	pit 筆 (pit)	brush pen
siën sen 先生 (xiānshēng)	sianseng 先生 (sian-seng)	Sir
sing 姓 (xìng)	sheⁿ 姓 (sèⁿ)	family name
soe 素 (sù)	tjiatjaj 食齋 (chia̍h-chai)	vegetarian
toeng tjia 東家 (dōng jiā)	tauwké, toke 頭家 (thâu-ke)	head of a business
wan seng 晚生 (wǎnshēng)	boan seng 晚生 (bóan-seng)	I (self-deprecatory)

稱·自 己 家 中 人
Panggilan dalam roemah sendiri poenja orang

稱 PANGGILAN	答 PENJAOETAN
祖父 tjoe voe engkong 祖母 tjoe moe oema	孫 soen tjoe tjoe
父親 voe tjing papa 母親 moe tjing mama	男 nan 兒子 el tje artinja anak
哥哥 ko ko engko 嫂嫂 saw saw enso	弟 ti (ade)

FIGURE 7.4   An explanation of kinship terms in *Hoa Woe Ho Pie*
CHUN FOO CHUN, *HOA WOE HO PIE* (BATAVIA: CHUN LIM &CO, 1920).
NO KNOWN COPYRIGHT HOLDERS

Most Mandarin words for the names of family members, too, lacked precise Malay equivalents. Fortunately, many Chinese-Indonesian families still used the corresponding Hokkien terms, even when their dominant language had shifted to Malay (Fig. 7.4). For the designation of maternal cousins and their children, for example, the Mandarin element *piaw* 表 (biǎo) corresponded perfectly to Hokkien *piaw* 表 (piáu). For siblings from the same paternal clan, Mandarin *thang* 堂 (táng) corresponded to Hokkien *tong* 堂 (tông).

Table 7.7 lists additional Mandarin kinship terms and their Hokkien/Sino-Malay translations, based on lists given in two of the textbooks (Chun 1920: 104–14; Shen-Tjia 1934: 245):

TABLE 7.7   Mandarin kinship terms and their Sino-Malay glosses

Mandarin	Sino-Malay	English gloss
ko ko 哥哥 (gēgē)	engkoh 俺哥 (ńg-ko)	older brother
koe moe 姑母 (gūmǔ)	kòh 姑 (ko·)	paternal aunt
koe tjang 姑丈 (gūzhàng)	kò-tio 姑丈 (ko·-tiōⁿ)	husband of paternal aunt
law pee 老伯 (lǎobó)	empee 俺伯 (ńg-peh)	older paternal uncle
law sioeng 老兄 (lǎoxiōng)	engkoh 俺哥 (ńg-ko)	older brother
law soe 老叔 (lǎoshū)	intjek 俺叔 (ńg-chek)	younger paternal uncle
pen tjia 本家 (běnjiā)	tjinlang 親人 (chhin-lâng)	distant relative
po foe 伯父 (bófù)	empe 俺伯 (ńg-peh)	older paternal uncle
saw saw 嫂嫂 (sǎosǎo)	enso 俺嫂 (ńg-só)	older brother's wife
shen moe 嬸母 (shěnmǔ)	entjim 俺嬸 (ńg-chím)	wife of younger paternal uncle
shoe foe 叔父 (shūfù)	entjek 俺叔 (ńg-chek)	younger paternal uncle
shoe moe 叔母 (shūmǔ)	entjim 俺嬸 (ńg-chím)	wife of younger paternal uncle
siong 兄 (xiōng)	engkoh 俺哥 (ńg-ko)	older brother
ta koe 大姑 (dàgū)	kòh 姑 (ko·)	older paternal aunt
ta yi 大姨 (dàyí)	toa-i 大姨 (tōa-î)	older maternal aunt
tang saw 堂嫂 (tángsǎo)	tong hia so 堂兄嫂 (tông-hiaⁿ-só)	wife of older brother of the same clan
tang sioeng 堂兄 (tángxiōng)	tong hia 堂兄 (tông-hiaⁿ)	older brother of the same clan
tang ti 堂弟 (tángdì)	tong te 堂弟 (tông-tē)	younger paternal male cousin
tang ti foe 堂弟婦 (tángdìfù)	tong te poe 堂弟婦 (tông-tē-pū)	younger paternal female cousin
tjêng tjoe foe 曾祖父 (zēngzǔfù)	kongtjo 公祖 (kong-chó·)	paternal great-grandfather
tjêng tjoe moe 曾祖母 (zēngzǔmǔ)	ma'tjo 媽祖 (má-chó·)	paternal great-grandmother
tjie tjie 姐姐 (jiějiě)	intji 俺姊 (ńg-chí)	older sister
tjioe foe 舅父 (jiù fù)	engkoe 俺舅 (ńg-kū)	maternal uncle
tjoe foe 祖父 (zǔfù)	engkong 俺公 (ńg-kong)	paternal grandfather
yi moe 姨母 (yímǔ)	ie 姨 (î)	maternal aunt
yi tjang 姨丈 (yízhàng)	ie-tio 姨丈 (î-tiōⁿ)	husband of maternal aunt

TABLE 7.8   Mandarin dictionary entries and their Sino-Malay glosses

Mandarin	Sino-Malay	English gloss
kan-lang 橄欖 (gǎnlǎn)	kana 橄欖 (kaⁿ-ná)	Chinese olive
kan-tso 甘草 (gāncǎo)	kamtjo 甘草 (kam-chhó)	liquorice root
ko-le-shen 高麗參 (gāolíshēn)	kolesom 高麗參 (ko-lê-som)	ginseng
kung-shi 公司 (gōngsī)	kongsie 公司 (kong-si)	company
kwa-tse 瓜子 (guāzǐ)	kwatji 瓜子 (koa-chí)	melon seeds
li 梨 (lí)	boeah laij (buah 'fruit' + 梨 lâi)	pear
liu-zun-yin 呂宋煙 (lǔsòng yān)	lisong 呂宋 (lī-sòng)	Luzon cigarettes
lung-yian 龍眼 (lóngyǎn)	lengkeng 龍眼 (lêng-kéng)	longan fruit
ma-feng 麻風 (máfēng)	taij-ko 癩哥 (thái-ko)	leprosy
mah-tsio 麻雀 (máquè)	mah-tjiok 麻雀 (mâ-chhiok)	mahjong
pian-sin-pu-soei 半身不遂 (bànshēn bùsuí)	piansoei 半遂 (piàn-sūi)	paralysis
sheng-jih 生日 (shēngrì)	she-djit 生日 (seⁿ-jit)	birthday
tau-yia 豆芽 (dòuyá)	tauge 豆芽 (tāu-gê)	bean sprouts
thaw 桃 (táo)	sianto 仙桃 (sian-thô)	peach
yang-mei 楊梅 (yángméi)	yangbwee 楊梅 (iâng-bôe)	syphilis

DATA TAKEN FROM LI (1931)

We see the same phenomenon in Li Joek Koey's 1931 Mandarin-Malay-Dutch-English dictionary, in which several "Malay" glosses of Mandarin entries are likewise Hokkien-derived (Table 7.8). In addition, a small number of glosses come from other Sinitic varieties, e.g. *pipah* 'loquat', *lobak* 'daikon', *sumok* 'sappanwood', and *tzifa* 'persimmon flower', respectively from Mandarin *pípá* 枇杷, Cantonese *lo⁴ baak⁶* 蘿蔔, Hakka *sû-muk* 蘇木, and Hakka *chhụ-fâ* 柿花.

In some textbooks, the Malay glosses reveal grammatical influence from the Chinese source text on account of their atypical idiom. The sentence *Tjé lie ti tjië tao hěn kan tjing* 這裡的街道很乾净 'The roads here are very clean', for example, is translated into Malay as *Disini poenja djalanan bersi sekali* (Shen-Tjia 1940: 114–15), rather than *Djalanan disini bersi sekali*. Chun Foo Chun in particular had a tendency to translate Chinese idioms literally into Malay. Thus, *khan soe* 看書 (kànshū) 'to read', *ming pê* 明白 (míngbai) 'to understand', *ta seng* 大聲 (dàshēng) 'loud', and *tjo tje* 坐車 (zuòchē) 'to take a car' become respectively *liat boekoe* 'to see books', *ngarti trang* 'to understand clearly', *swara besar* 'big voice', and *doedoek kereta* 'to sit in a car'. For the verbal constructions

tse jen 喫煙 (chīyān) 'to smoke', *tse tjah* 喫茶 (chīchá) 'to have tea', and *tse tjioe* 喫酒 (chījiǔ) 'to have alcohol', Chun's Malay translations faithfully display *makan roko* 'to eat cigarettes', *makan thee* 'to eat tea', and *makan arak* 'to eat alcohol', even though the word *minoem* 'to drink' would have been more natural. In Mandarin itself, the usage of the verb for 'to eat' in such constructions is currently considered as "southern non-standard", possibly reflecting Chun's Hakka origins. Furthermore, the copulative verb *sze* 是 (shì) and the adjectival particle *hen* 很 (hěn) are consistently translated into Malay as respectively *betoel* 'indeed' and *amat* 'very', yielding awkward translations as both words can (and should) be omitted in idiomatic Malay. In other cases, Chun translates Mandarin terms with a short Malay description, e.g. *joan toeng* 圓通 (yuántōng) 'accommodating' as *tida soeka kasi orang marah* 'does not like to upset people' and *tjhian koeng* 謙恭 (qiāngōng) 'humble' as *tida mata tingi kaja miskin sama rata* 'is neither haughty nor distinguishes between rich and poor'.

5 Phrasing Chineseness

Having investigated the linguistic peculiarities of both the Mandarin source language and the Sino-Malay target language, we now take a closer look at the contents of the textbooks. Language manuals and phrasebooks generally provide miscellaneous examples of the way power hierarchies and other social information are linguistically encoded. A plethora of European-authored Malay phrasebooks, for example, entered the markets of the Netherlands Indies and British Malaya from the late-nineteenth century, promising the prospective colonial servant, soldier, or tourist efficient ways to communicate with the "native population". This topic deserves a separate study and could fruitfully be connected to present-day Singaporean guidebooks on talking to (Indonesian and other) domestic workers. Malay phrasebooks intended for Chinese readers, too, have been discussed in more detail elsewhere (Hoogervorst 2021). In what follows, I will limit myself to the contents of Mandarin textbooks in Malay.

The *Classified Conversations*, which served to prepare Indies-born Chinese for a trip to China, contains many example sentences that take place in trams, trains, and boats. The phrasebook clearly aimed to portray the newly established ROC as a modern society. No better illustration can be given than the following selected sentences, given in the context of a Western-style dance party (Chu et al. 1912: 64–65):

要我扶麼	Yao wo foe mo?	Shall I offer you my arm?	Bolehkah saja bri tangan saja padamoe? (toentoen)
那些太太穿的衣服很雅致很清潔[a]	Na see tai tai tjoan tih i foeh hen ya tzi hen tjiêh tjin	The ladies are very elegantly and neatly dressed	Njonja njonja berpake amat baik dan rapi
太太，我可以和你作方形跳舞麼	Tai tai, wo koo i hoo ni tjo fang sjing tiao woe mo?	May I have the pleasure of dancing this quadrille with you, Madam?	Bole saja plesir dansa di ini dansaan sama kau njonja?
跳舞叫我頭暈	Tiao woe tjhiao wo thoouw yuen	Walzing [sic!] makes me giddy	Dansa bikin saja sempojongan
我給你一杯檸檬水	Wo kie ni i pei lin-mon soei	Allow me to offer you a glass of lemonade	Kasi saja permisi boeat soegoeken kau satoe glas limonade

a 清潔 (qīngjié) for *tjiêh tjin* should presumably have been 潔淨 (jiéjìng) in view of the pronunciation given. Both words could be translated with English 'clean' and/or Malay *rapi*.

Chun's *Hoa Woe Ho Pie*, with its focus on the Netherlands Indies, contains various example sentences in a classroom setting (Chun 1920: 134–37). These are clearly inspired by contemporaneous schoolbooks from China. A school teacher's explanation of the aphorism *Sze noeng koeng sang tjiay ih tang ping* 士，農，工，商，皆宜當兵 'scholars, farmers, laborers, and businessmen should all become soldiers', for example, can be traced back to a popular schoolbook from the ROC titled *Mandarin Textbook* 國語教科書, which was also used by Tiong Hoa Hwe Koan schools.[30] Other example sentences in *Hoa Woe Ho Pie* confirm that a great nation should possess a great army. However, as the writer reassured, 'Java is a territory of the Netherlands': *Tjauw wa sze ho lan kwô tî sôe ti* 爪哇，是荷蘭國的屬地 (Chun 1920: 148). Notwithstanding Holland's political hegemony, however, the textbook left little doubt that the loyalty of the colony's Chinese population belonged to the ROC. The following questions-and-answers illustrate this point (Chun 1920: 120–21):[31]

---

30   Copies from the Netherlands Indies can be found at the Leiden University Library (Sinol. 153236). For more context, see Hou (2017).

31   Here and below, the translations into English are mine.

中國人，不讀中國書, 好麼?	Tjoeng kwo jin poe toe tjoeng kwo soe hao mo?	orang tjina tidak sekolah tjina baik tida?	'Is it a good thing if Chinese people do not study China?'
很不好，很羞愧	Hen pu hao, hen sioe khwe	amat tida baik amat maloe	'It's very bad, very embarrassing'
怎麼樣, 很羞愧?	Tjan me jang hen sioe khwe?	bagi mana amat maloe?	'Why is it very embarrassing?'
中國人，不讀中國書，不曉得中國字，很羞愧	Tjoeng kwo jin poe toe tjoeng kwo soe poe siaw te tjoeng kwo tze hen sioe khwe	orang tjina tida sekola tjina tida taoe soerat tjina amat maloe	'Chinese people who do not study China and cannot write Chinese are very embarrassing'
不讀中國書，不曉得講中國話，天天講馬來話，更羞愧	Poe toe tjoeng kwo soe, poe siaw te tjiang tjoeng kwo wa, thin thin tjiang ma le wa, keng sioe khwe	tida sekola tjina tida taoe bitjara tjina hari hari bitjara malajoe lagi maloe	'Not studying China, not being able to speak Chinese, and speaking Malay every day is even more embarrassing'
你愛中國麼?	Ni ay tjoeng kwo mo?	loe tjinta negeri tjina tida?	'Do you love China or not?'
我是中國人，我很愛中國	Wo sze tjoeng kwo jin, wo hen ay tjoeng kwo	saia djadi orang tjina saia amat tjinta negeri tjina	'I'm Chinese, I love China very much'
你將來要去中國麼	Ni tjiang le yaw tjoey tjoeng kwo mo?	loe nanti maoe pegi negeri tjina tida?	'Do you want to go to China in the future?'
我一定要去中國	Wo ie ting yaw tjoey tjoeng kwo	saia misti maoe pegi negeri tjina	'I definitely want to go to China.'

In the second volume of *Kuo-Yü Zonder Goeroe*, we find a very similar exchange, shaming Indies-born Chinese who cannot speak Mandarin (Shen-Tjia 1940: 9, 15–16). A selection of sentences is given below:

您說得那末準確，真可佩服	nien shuo tê na mo tj'oen tj'io, tjên k'ê p'ei toe	Saia merasa kagoem sekali jang toean bisa bitjara begitoe teges	'You speak so accurately, it's very admirable'
不敢當，聽說荷印的華僑，除了由祖國來的，很少能說中國話，這是確實嗎	poe kan tang, t'ing shuo ho yin ti hua ch'iao, tj'oe liao yoe tsoe kuo lai ti, hên shao nêng shuo tjoeng kuo hoa, tjé shih tj'io shih ma?	Trima kasi. Saia denger Hua Kiao di Hindia Blanda, selainnja totok, sedikit sekali jang bisa bitjara Tionghoa. Apatah ini betoel?	'You flatter me. I hear that Overseas Chinese in the Netherlands Indies, except those born in China, are rarely able to speak Chinese. Is this true?'
是確實的，大概有一半不懂自己的語言，這實在是慚愧得很	Shih tj'io shih ti, ta kai yoe i pan poe toeng tzŭ tji ti yü yan, tjé shih tjai shih tj'an k'oei tê hên.	Ini memang betoel, kira-kira ada separo jang tida bisa bahasanja sendiri. Ini sebetoelnja berasa maloe sekali	'This is true, about half of them don't understand their own language. This is really very embarrassing.'
中國人不會說中國話，當然是害羞的. 而且會給外國人看不起	Tjoeng kuo rĕn poe hoei shuo tjoeng kuo hoa, tangran shih hai sioe ti. Ěrl tj'iĕ hoei kei wai kuo rĕn k'an poe tj'i	Bangsa Tionghoa tida bisa bitjara bahasanja tentoe sekali maloe, dan lagi dipandang renda oleh lain bangsa.	'Chinese people who cannot speak Chinese are definitely very embarrassing, and will also be looked down on by foreigners.'
因為有了這種感覺，所以我家兄弟姐妹都趕快學國語了	Yin wei yoe liao tjé tjoen kan tjio, so i wo tjia hsioeng ti tjie mei toe kan k'oai hsuën kuo yu liao.	Oleh karna ada ini pengrasahan, maka saia poenja soedara lelaki dan prampoean semoea pada boeroe-boeroe berladjar kuo-yu.	'Because of this feeling, my brothers and sisters are quickly learning Mandarin.'

Whereas the first volume of *Kuo-Yü Zonder Goeroe* contained very few example sentences of a distinct ideological nature, the second volume insisted at various sections that Indies-born Chinese should learn about China and preferably spend time there. Like the *Classified Conversations* almost three decades earlier, the book's conversations took place in the cities, trains, and boats of China.

Both books featured conversations on seasickness and ways to handle it, for example. On the topic of dancing, however, Mrs. The proved more conservative than her predecessors. While the latter treated the topic with the awe of novelty, the following conversation in *Kuo-Yü Zonder Goeroe* reveals that its author had little positive to say about the encroachment of Western dancing on the Chinese culture (Shen-Tjia 1940: 237):

我想不到，現在有這麼多中國人跳舞	Wo siang poe tao, hsiën tjai yoe tjé mo to tjoeng kuo rĕn t'iao woe.	Saia tida sangka sekarang banjak orang Tionghoa bisa dansa	'I didn't expect there would be so many Chinese people these days who dance'
這就是所謂時髦的中國人，難道你反對中國人跳舞嗎	Tjé tjioe shih so wei shih mao ti tjoeng kuo rĕn, nan tao ni fan toei tjoeng kuo rĕn t'iao woe ma?	Ini jalah jang dibilang orang Tionghoa modern. Masatah kaoe anti orang Tionghoa dansa?	'These are the so-called modern Chinese. Are you against Chinese people who dance?'
絕對反對. 我們出去再繞個圈子吧	Tj'uëh toei fan toei. Wo mĕn tj'oe tj'u tjai rao kĕ tj'uan tzŭ pa!	Tentoe anti. Mari kita kloear moeter lagi sekali	'I'm absolutely against it. Let's go out and walk around some more.'

6 Concluding Remarks

The enduring popularity of Mandarin textbooks in Chinese-Indonesian circles negate Oudendijk's observations made in 1913 that "those who have settled here permanently will probably realize that the Mandarin dialect has little 'market value'" and that "while the written Chinese language can still be counted among the sciences that are useful for the (Chinese) subjects in the Netherlands Indies society, the spoken North Chinese (Mandarin) dialect cannot be considered for this at all".[32] At the same time, the fact that few of these

32 [...] zullen zij die hier voor goed gevestigd zijn wel inzien dat het Mandarijnsche dialect weinig "marktwaarde" heeft (van der Wal 1963: 258); [...] kan men de geschreven Chineesche taal nog rekenen tot die wetenschappen die voor onze (Chineesche) onderdanen in de Nederlandsch-Indische samenleving nuttig zijn, het gesproken Noord Chineesche (Mandarijnsche) dialect kan daarvoor heelemaal niet in aanmerking gebracht worden (ibid. 263).

educational books – oft-reprinted and well-advertised as they were – can be found in more than one library testifies to a later disinterest in the material.[33] This is unfortunate, as Chinese-Indonesian linguistic history can only be fully understood by taking into account all elements in the equation: Malay, Mandarin, Hokkien, Javanese, and even Dutch.

In this chapter, hence, I have tried to demonstrate that several insights of historical and linguistic relevance can be gleaned from the pages of Mandarin educational texts, as is the case with phrasebooks and textbooks more generally. As regards the romanization of Mandarin, we may safely conclude that no available system fully "worked"; rather than a consistently implemented system, what arose was a fluid set of conventions. Testimony to this fluidity is the fact that some authors adopted a different romanization over time, while others openly solicited their readers' preferences in this arena. I have focused more on lexicon and textual contents than grammar, which is comparatively difficult to connect with broader social developments, yet the material will undoubtedly also be relevant to scholars interested in the syntactical development of Mandarin.

The latest primary source examined for this chapter dates from 1945. Mandarin-related materials from the 1950s presumably reveal quite different dynamics. Amidst the triumph of the Indonesian nation-state, books promoting loyalty to another country became politically sensitive. Nevertheless, both the PRC and the ROC – the latter now operating from Taiwan – continued to assert their influence on Chinese-Indonesians in subtle, officially approved ways (Suryadinata 1972; Sai 2006; Zhou 2019). This came with linguistic implications too. In 1955, the PRC adopted a standardization closer to the colloquial language of Beijing – known as Common Speech or *Pǔtōnghuà* 普通話 – while Taiwan retained the pre-existing standardization that, though also based on the Beijing dialect, preserved a layer of Nanjing lexical influence associated with intellectual speech (Li 2004: 103; Simmons 2017: 65–66). The enforcement in the PRC of simplified characters and of the *Pīnyīn* 拼音 romanization, too, took an upward turn from the 1950s. It stands to reason to assume that pro-Beijing and pro-Taiwan schools taught a slightly different form of Mandarin in 1950s Indonesia, until the latter were banned in 1958 (cf. Suryadinata 1972: 67–70). In precisely these years, calls to outlaw Chinese education and Chinese newspapers rang increasingly louder (cf. Elson 2008: 176–77). Just when only

---

33   This is especially the case for textbooks written in Malay. The situation for textbooks in Chinese, a collection of which is kept at the Wang Gungwu Library in Singapore, is considerably more favourable.

one type of Mandarin was left for Chinese-Indonesians to choose from, their very identity had become illicit.

The Mandarin currently taught in Indonesia is chiefly *Pǔtōnghuà*, in simplified characters. It is avidly studied by Chinese-descended as well as "indigenous" Indonesians. In addition, numerous Indonesian migrant workers have gained fluency – but rarely literacy – in Taiwanese Mandarin. This variety is also taught in a small number of schools. While knowledge of Mandarin is clearly on the rise in Indonesia and elsewhere in Southeast Asia, proficiency in Hokkien, Hakka, Cantonese and Teochew is declining. Indeed, the ongoing Chinese Revival leaves little room for Sinitic varieties other than Mandarin (Sai 2010; Setijadi 2015; Stenberg 2015). As I have demonstrated, this project of homogenization was kick-started a century earlier. Whether it will lead to new episodes of shaming Chinese-Indonesians deemed not fluent enough remains an open question. If anything, doing so would undermine a fascinating history of plurilingualism and fluctuating proficiencies, in which linguistic practices were deeply intertwined with the political pressures of the day.

Acknowledgements

References

China Weekly Review, Who's Who in China: Biographies of Chinese Leaders (Shanghai: China Weekly Review, 1936).

Chu, K.S., Kim S.L. and Lim B.S., *The Chinese, English and Malay Classified Conversations* (Batavia: Chun Lim &Co, 1912).

Chun Foo Chun, *Homaro: Hatoeran Oentoek Memake aken Roepa-Roepa Obat atawa Penoendjoek Roepa-Roepa Penjakit dan Obatnja* (Batavia: Chun Lim &Co, 1915).

Chun Foo Chun, *S.A.N.O.P.H.I.: Segala apa njang Orang Prampoean Haroes Ingat* (Batavia: Lim Bok Sioe, 1917).

Chun Foo Chun, *Hoa Woe Ho Pie* (Batavia: Chun Lim &Co, 1920).

Chun Foo Chun, *Sin Woe Hoa Tze Tian: Woordenboek Melajoe en Tiong-hoa* (Batavia: Chun Lim &Co, 1930).

Elson, R.E., *The Idea of Indonesia: A History* (Cambridge: Cambridge University Press, 2008).

Goh, Yeng-Song, *Teaching Chinese as an International Language: A Singapore Perspective* (Cambridge: Cambridge University Press, 2017).

Hau Lai Ying 侯勵英, '新學制下的兒童教科書: 民初《 國語教科書》 的知識傳播', in Lee Fan (李帆), Hon Tze-ki (韓子奇) and Au Chi Kin (區志堅) (eds.), 知識與認同: 現代學者論教育及教科書 (Hong Kong: Chung Hwa Book Company, 2017): 225–53.

Heryanto, Ariel, 'Ethnic Identities and Erasure: Chinese Indonesians in Public Culture', in Joel S. Kahn (ed.), *Southeast Asian Identities: Culture and the Politics of Representation in Indonesia, Malaysia Singapore, and Thailand* (London and New York: I.B. Tauris, 1998): 95–114.

Hoogervorst, Tom G., 2021. *Language Ungoverned: Indonesia's Chinese Print Entrepreneurs, 1911–1949.* Ithaca, NY: Cornell University Press, 2021.

Iskan, Dahlan, 'Wajah Baru Koran Si Oom', *Tempo* 12/18 (1982): 28.

Knoetsen, John F., Tan Giem Hok and M.B. Sastro Widjojo, *Java-Tourists: How to Speak Malay, Javanese and Sundanese* (Batavia: Tjiong Koen Bie, 1911).

Knoetsen, John F., *East Asia* (2nd edn., Batavia: Tjiong Koen Bie, 1912 [1910]).

Knoetsen, John F., *Handleiding om Chineesche Karakters te leeren schrijven en lezen* (Batavia: De "Pertoendjoengan", 1931).

Knoetsen, John F., *Japansche taal voor Hollandsch en Maleisch sprekende toeristen, Hollandsch-Japansch-Maleisch-Engelsch* (Batavia: De "Pertoendjoengan", 1934).

Knoetsen, John F., *De Procureur* (Batavia: De "Pertoendjoengan", 1935).

Kuiper, Koos, *The Early Dutch Sinologists (1854–1900): Training in Holland and China, Functions in the Netherlands Indies* (Leiden: Brill, 2017).

Kwartanada, Didi, 'Mandarin Comes to the South Seas: The Making of Chinese Education in Early Twentieth Century Java', *Asian Culture* 42 (2018): 37–54.

Li, Chris Wen-Chao, 'Conflicting notions of language purity: The interplay of archaising, ethnographic, reformist, elitist and xenophobic purism in the perception of Standard Chinese', *Language & Communication*, 24 (2004): 97–133.

Li Joek Koey, *A Classified Dictionary of Chinese-Malay-Dutch-English with Chinese pronunciation* (Batavia: Kuo Min Book Co., 1931).

Mahdi, Waruno, 'Linguistic variety in later nineteenth-century Dutch-edited Malay publications', *Nusa*, 60 (2016): 107–85.

Moerman, J., *In en om de Chineesche Kamp* (Batavia: G. Kolff, 1929).

Morgan, Evan, *Chinese new terms: Revised and enlarged. With English translations, classifications, introduction and index* (Shanghai: Kelly & Walsh, 1932).

Ng, Kim Chew, 'Sinophone/Chinese: "The South where Language is lost" and Reinvented, translated by Brian Bernards', in Shu-mei Shih, Chien-hsin Tsai and Brian Bernards (eds.) *Sinophone Studies: A Critical Reader* (New York: Columbia University Press, 2013): 74–92.

Nio Joe Lan, 'De omgangstaal der Chineezen in Indonesië', *De Indische Gids*, 54/2 (1932): 1089–94.

Nio Joe Lan, 'Iets over Chineesche namen en over het Chineesch in Indonesië', *De Indische Gids*, 55/1 (1933): 410–18.

Oetomo, Dédé, 'The Chinese of Indonesia and the Development of the Indonesian Language', *Indonesia* (1991, special issue): 53–66.

Oey Tjiang Hoay, *Kitab Logat Baroe Tionghoa-Melajoe* (Shanghai: Modern Languages Association, 1935).

Sai Siew-Min, *Representing the Past of Chinese Language Education: Language, History and Chinese Identities in Indonesia* (PhD thesis, University of Michigan, 2006).

Sai Siew-Min, 'Pugilists from the Mountains: History, Memory, and the Making of the Chinese-Educated Generation in Post-1998 Indonesia', *Indonesia*, 89 (2010): 149–78.

Sai Siew-Min, 'Mandarin lessons: Modernity, colonialism and Chinese cultural nationalism in the Dutch East Indies, c.1900s', *Inter-Asia Cultural Studies* 17/3 (2016): 375–94.

Setijadi, Charlotte, 'Being Chinese Again: Learning Mandarin in Post-Suharto Indonesia', in Li Wei (ed.), *Multilingualism in the Chinese Diaspora Worldwide: Transnational Connections and Local Social Realities* (London: Routledge, 2015): 141–57.

Shen-Tjia, Ms. The Chung, *Kuo-Yü Zonder Goeroe* (Surabaya: New China, 1934).

Shen-Tjia, Ms. The Chung, *Kuo-Yü Zonder Goeroe: Djilid Doea* (Surabaya: New China, 1940).

Shen-Tjia, Ms. The Chung, *Boekoe Beladjar Alphabet Kuo-Yin* (Surabaya: New China, 1942).

Simmons, Richard Vanness, 'Whence Came Mandarin? Qīng Guānhuà, the Běijīng Dialect, and the National Language Standard in Early Republican China', *Journal of the American Oriental Society* 137/1 (2017): 63–88.

South Coblin, W., 'A Diachronic Study Of Míng Guānhuà Phonology', *Monumenta Serica* 48/1 (2000): 267–335.

South Coblin, W., *Modern Chinese phonology: From Guānhuà to Mandarin* (Paris: EHESS, 2007).

Suryadinata, Leo, 'Indonesian Chinese Education: Past and Present', *Indonesia* 14 (1972): 49–71.

Tip Top Toe, *Gadis jang dibawa lari* (Batavia: Sin Po, 1923).

Tshen, Eric H., Oey Tjiang Hoay and Lim Bok Sioe, *Han Woe Tzi Tian: Kitab Lograt Tiong Hoa-Melajoe* (Batavia: Kho Tjeng Bie, 1915).

van der Wal, S.L., *Het Onderwijsbeleid in Nederlands-Indië 1900–1940: Een Bronnenpublikatie* (Groningen: J.B. Wolters, 1963).

Wang, Anyi, 'The Fate of Language (Parts 1 and 2)', *Xingzhou ribao, xingyun ban* [Sin Chew daily starcloud supplement] (13 July and 7 December 1993).

Wangkar, Max, 'Suratkabar Jawa Pos adalah Dahlan Iskan: Bagaimana dalam 20 tahun satu koran kecil jadi dominan di Indonesia', *Pantau* 2/13 (2001): 12–25.

Weng, Jeffrey, 'What Is Mandarin? The Social Project of Language Standardization in Early Republican China', *The Journal of Asian Studies* 77/3 (2018): 611–33.

Yoe Wan Fei, *Lim's Boekoe Peladjaran Kuo-Yü Zonder Goeroe* (Jakarta: Lim's Boekhandel, 1945).

Zhou Taomo, *Migration in the Time of Revolution: China, Indonesia, and the Cold War* (Ithaca: Cornell University Press, 2019).

Conclusion

Southeast Asia's Sinitic "voices", in all their heterogeneity, are most easily perceived through a comparative analysis. It has been a true pleasure to capture in one volume the perspectives of so diverse a selection of active speakers, passive speakers, rememberers, academics, and consultants. Their stories of language contact and linguistic change, whether based on original fieldwork, archival research, or a combination of both, reveal the Sinophone in new and unexpected ways. No less exciting is the incorporation of rarely quoted sources in Chinese and Southeast Asian languages. In this volume, we have attempted to expand the field of Sinophone Studies, infuse it with texts that rethink the limits of its original scope, and provide some directions for future scholarship. When taken together, the chapters of this book lay some of the groundwork to diversify the Sinophone conceptually, regionally, and in disciplinary terms. In particular, we have advocated for the appreciation of regional languages (方言 *fangyan*) as significant strands within Sinophone Studies, complementing earlier works that prioritized the analysis of literature and cinema.

Language sits at the core of the scholarly intervention we propose. A focus on Southeast Asia, on account of its diversity and long contact history with the Sinitic heartlands, is critical to this endeavour. To appreciate the significance of language change, lexical borrowing, creolization, and plurilingualism seen among Chinese-descended Southeast Asians, we have insisted on examining the region's Sinitic and Sinitic-influenced languages in their historical, social, and cultural contexts. The individual chapters provide concrete instances of the insights thus obtained. In general, the Sinitic languages from China's southern provinces that made their way into the Nanyang have become intriguingly locale-specific over time. This has been shown in detail for Timor Hakka (Chapter 2), Cambodian Teochew (Chapter 3), and Penang Hokkien (Chapter 5), and is the case for numerous other varieties. Shaped by their new geographies and prolonged interaction with (other) local languages, Cambodian Teochew now differs significantly from Singapore Teochew, Penang Hokkien from Burmese Hokkien, and Timor Hakka from Malaysian Hakka. These varieties – many of which are critically endangered – tell intimate stories of regions, communities, and even families.

While the chapters of this volume differ in the specific voices and accents they foreground, they are held together by a number of connections. They all underline plurilingualism as a dominant force running through Southeast Asia's Sinitic landscapes. They also instantiate the counterpull between low-status

regional vernaculars – both Sinitic and Sinitic-influenced – and languages of high prestige, such as Mandarin and the official languages of Southeast Asia's nation-states. To widen the range of voices encapsulated in this book, we have provided space for the hybrid, the creolized, and the marginalized. Due in part to their resistance to standardization, the region's "non-Mandarin" Sinitic varieties have incorporated multiple tongues, including other Sinitic, local Southeast Asian, and European languages. In addition to lexical and grammatical convergence, many Southeast Asian cities – where people from different Sinitic backgrounds were clustered together – also became sites of koineization; the mixing of dialectal features.

Some of the vernaculars emanating from these contact histories assumed important communicative needs in Southeast Asia and beyond. These were typically Sinitic-influenced rather than Sinitic, such as Chinese Pidgin English and Bazaar Malay. Such predominantly spoken languages should be envisioned, simultaneously, as the products and catalysts of language contact and the vectors and donors of new vocabulary. Chinese Pidgin English, for example, offers concrete instances of the multidirectional lexical trajectories underpinning eighteenth- and nineteenth-century Sino-European trade, exemplified by such words as *cumshaw* 'present, gratuity' from Hokkien *kam sia* 感謝 'thank you', *fytie* 'hurry' from Cantonese *fai di* 快啲, and *savvy* or *sha-pi* 'to understand' from Portuguese *saber* (Chapter 1).

This volume's pan-Southeast Asian scope has allowed for comparative insights. A number of commonalities show up across the region. These include broadly shared words, such as 唐人 (*tng lang, tong yan, tangren*, etc.) 'Chinese person', 舵公 (*tai kong, to gung, duogong*, etc.) 'helmsman', 頭家 (*thau ke, tau gaa, toujia*, etc.) 'boss', 公司 (*kong si, gung si, gongsi*, etc.) 'company', and various other examples in the realms of business, cuisine, and culture. Even families in which Sinitic proficiency has been lost over the generations – a common process throughout Sinophone Southeast Asia – typically retain "heritage words" from their ancestral varieties, including kinship terms. This is common among the descendants of migrants in general. Beyond such retentions, Sinophone Southeast Asia has also become an incubator of lexical innovation. Its wealth of creativity can be seen in the designation of novel concepts – from tin mines and modes of transportation to balloons and mobile phones – and specific food items. The jicama (*Pachyrhizus erosus*), enjoyed throughout East and Southeast Asia, provides a case in point. This New World import, introduced into the Sinosphere in early-modern times, gave rise to miscellaneous local nomenclature including *tau aw jiu* 豆仔薯 (Taiwan), *kua jiu* 芥薯 (Myanmar), *sa koh* 沙葛, and *bang-kuang* 芒光 (Singapore and Malaysia) (Chapter 4).

These examples are from Hokkien, but counterparts can be found in many other Sinitic languages.

The parallel linguistic developments of Sinophone Southeast Asia can be fruitfully analysed through word histories. All localized Sinitic varieties acquired new words for previously unfamiliar concepts (expansive borrowing), in addition to concepts for which the corresponding inherited word was lost – if not deemed inadequate – over the generations (replacive borrowing). This explains, among many other things, the presence of loanwords from Portuguese, Indonesian, and Tetun in Timor Hakka, including *kapoŋ* 'bonnet, hood' from Portuguese *capô* or Tetun *kapó* and *pasat* 'market' from Indonesian *pasar* or Tetun *basar* (Chapter 2). Another way of designating new concepts was through processes of compounding and semantic extension. Some speakers of Cambodian Teochew, for example, use the locally invented term *tek²⁻⁵kɔŋ³³kue⁵²* 竹管粿 'bamboo cake' to refer to the local snack *kralan* ក្រឡាន (Chapter 3). Penang Hokkien in particular is replete with colourful expressions – including *tāu-gê·-jī* 豆芽字 (lit. 'bean sprout letters') for the Arabic-derived Malay script (*Jawi*) and *chiáh-Kè-lêng-á-pūiⁿ* 食吉寧仔飯 (lit. 'eating Indian curries') for spending time in prison (Chapter 5) – but this is a feature shared by Southeast Asia's Sinitic languages more broadly.

In the opposite direction, most of the larger Southeast Asian languages have been enriched by loanwords from regional Sinitic varieties, which have generally become marginalized in recent times. The study of Cambodian Teochew (Chapter 3), for example, illustrates how the prolonged presence of Teochew speakers in Cambodia has left a linguistic imprint on Khmer, the official language, through such borrowings as Khmer (ʔaa) koŋ កុង 'grandfather' from Teochew (ʔa³³)koŋ³³ (阿) 公 and *sinsaae* ស៊ិនសែ 'doctor; teacher' from *siŋ³³ sɛ̃³³* 先生. Other national languages that have been enriched in this way include Thai, Indonesian, Tagalog, and Burmese, which display many of the same loanwords.

In addition to these instances of direct borrowing, new concepts were often designated through loan translations, in which the elements of local Southeast Asian languages were rendered word for word in the recipient Sinitic languages. The Khmer word *pʰlaɛɓɔ* ផ្លែប៊ើ (lit. 'butter fruit') for 'avocado', for example, is literally translated by some speakers of Cambodian Teochew as *gu⁵⁵⁻¹¹iu⁵⁵⁻¹¹kue⁵²* 牛油果, although other terms are used as well (Chapter 3). Along similar lines, Penang Hokkien exhibits the idioms *chiáh-hong* 食風 'to take the air; to take a holiday' from Malay *makan angin* (lit. 'to eat the wind'), *jíp-Hoan* 入番 'to convert to Islam' from *masuk Melayu* (lit. 'to enter Malaydom'), *kim-chio* 金蕉 'a kind of banana' from *pisang mas* (lit. 'gold banana'), *o·-iû-lâng*

烏油儂 'a type of Malay ghost' from *orang minyak* (lit. 'oil man'), and possibly *pē·ⁿ-chhù* 病厝 'hospital' from *rumah sakit* (lit. 'house of the ill') (Chapter 5).

An equally insightful category are loan blends, which combine a Sinitic and a non-Sinitic element. Such constructions are quintessentially Sino-Southeast Asian. We may recall the Cambodian Teochew expression *hɔu¹¹ka¹¹dɔu¹¹* 'to give someone a gift', consisting of Teochew *hɔu¹¹* 互 'to give' and Khmer *kadoʊ* កដូ 'gift', the latter being itself a French loanword (Chapter 3). Along similar lines, the Timor Hakka word for coffee, *kopi tʃha* 'coffee' (lit. 'coffee tea'), consists of Indonesian *kopi* 'coffee' and Hakka *chhà* 茶 'tea' (Chapter 2). Many more examples can be found across Southeast Asia's Sinitic varieties. These hybrid expressions also feature in Sinitic-influenced Malay varieties. Baba Malay, for example, has *buat suay* 'to cause misfortune', consisting of Malay *buat* 'to make' and Hokkien *suay* 衰 'misfortune' (Chapter 4). Java's Sino-Malay vernacular, at least in the 1930s, exhibited *boeah laij* 'pear', consisting of Malay *buah* 'fruit' and Hokkien 梨 *lâi* 'pear' (Chapter 7).

As can be expected, lexical borrowing chiefly took place in the domain of cultural vocabulary. The transmission of function words is relatively rare, but we do encounter some examples in Southeast Asia. Cambodian Teochew *pi³³* 'from', for example, goes back to Khmer *pi*: ពី in the same meaning. In many Sinitic varieties spoken in Indonesia, Malaysia, and Singapore, the word *tapi* 'but' is used in addition to – or instead of – inherited equivalents, including by the graduates of Mandarin-medium schools. In some Sinitic-influenced languages, by contrast, we find local words rearranged in a Sinitic pattern. Baba Malay and Bazaar Malay, for example, display such constructions as *ari satu* 'Monday', *dia punya* 'his; her', and *diaorang* 'they', which are modelled, respectively, after Hokkien *pai-it* 拜一, *i e* 伊的, and *i-lang ~ i-nang* 伊儂 (Chapter 4). The fact that these examples have found their way into other Malay varieties indicates a mainstreaming of what were historically pidgin features.

Another area for theoretical expansion lies within the notion of textuality itself. While the analysis of literary works is relatively commonplace in Sinophone Studies, we have called attention to texts of a more vernacular nature – encompassing the domains of theatre, popular printing, and language learning – as equally legitimate sources underpinning the rich and varied Sinitic and Sinitic-influenced landscapes of Southeast Asia. Interconnections between spoken and written language came to the fore in multiple chapters, raising questions about the extent to which this binary is tenable in the first place.

Upon comparing the volume's chapters, a number of observations can be made regarding the graphic variation of Sinophone Southeast Asia, which provides rich illustrations of Sinitic writing beyond Mandarin and other

more-or-less standardized varieties. Amidst this heterogeneity, two broader phenomena vie for attention: 1) competing systems of romanization, and 2) the use of Chinese characters (Sinographs) to transcribe colloquial and/or non-Sinitic words.

In the first category, we may call attention to romanizations based on non-Sinitic orthographies, such as the Indies-style romanization of Mandarin words like *tjha* (茶 *chá*) 'tea' and *koe moe* (姑母 *gūmǔ*) 'paternal aunt' based on Dutch (Chapter 7) or the Philippine-style transcription of Hokkien *di* 'you' (你 *li*), *kuay* 'quick' (快 *khuai*), and *lay* 'come' (來 *lai*) as seen in the playscripts of a prolific theatre leader (Chapter 6). Romanization adds layers of analysis to Sinophone texts, in particular those of a more hybrid, localized type. The term *tjoe-têng siao-hsüeh* for 'Chinese elementary school', historically attested among the Indies Chinese, provides a case in point: only its romanized form reveals that it is a combination of Hokkien *chiūⁿ-tŷg* 上堂 'to attend class' and Mandarin *xiǎoxué* 小學 'primary school' (Chapter 7). In historical Sinitic texts, too, the script typically obscures rather than illuminates the role of specific regional varieties. It is chiefly due to romanization – including by non-Chinese people – that historically important information about the provenance of words and names can occasionally be extracted (Chapter 1).

In the second category, we find Chinese characters that have defied standardization. These include Sinographs used to transcribe Southeast Asian or European words, but also inherited words from Sinitic varieties that historically lacked standardized characters. We have seen, for instance, variation in Southern Min varieties between the characters 予, 乎, and 互 for 'to give'. The Hokkien word *Kaoka*, a specific type of theatre, would be another example. Next to its most common characters 高甲 – with competing etymologies ranging from 'standing on the high stage and wearing armour' 登高臺穿盔甲 to 'high quality and A-grade' 高等甲等 – we find such alternative choices as 九甲, 九角, and 戈甲. Non-standard orthography features prominently in the handwritten playscripts used in this genre, as seen from the characters 戔 (*tsiàn*) 'war' 边 (*pian*) 'side, border', and 闩 (*kuan*) 'door, gate' corresponding to "standard" 戰 (*zhàn*), 邊 (*biān*), and 關 (*guān*) (Chapter 6). Penang Hokkien exhibits numerous additional examples, including *koa-sa-jī* 啩吵字 'a will or document of power of attorney' – in which *koa-sa* 啩吵 is a phonetic representation of Malay *kuasa* 'power' – and *chhiū-leng-pa* 樹朧芭 'rubber plantation', featuring *pa* 芭 from Thai *paa* ป่า 'forest' and the locally coined character 朧 (*leng*) for 'rubber' (Chapter 5).

Here, again, a combination of the spoken and written word is essential to arrive at a complete picture. An exclusive focus on the script may obscure regional or non-Sinitic influences, especially if the words in question are

written in characters selected to approximate their sound and meaning (phono-semantic matching). The regional origins of 紅毛丹 'rambutan' (lit. 'red-haired crimson') and 老君 'doctor' (lit. 'elderly lord'), for example, would be ambiguous without juxtaposing their Hokkien pronunciations and Malay etyma: *âng-mô·-tan* vs. *rambutan* and *ló-kun* vs. *dukun* (Chapter 5). This is particularly relevant since some concepts, including the word for 'rambutan', have entered the broader Sinosphere as graphic loans (Chapter 1). These character-based borrowings remind us to not lose sight of writing either. The rich Chinese textual record often provides valuable information on their distribution, popularity, and time depth. Sinographs are also crucial to understand various primary sources on Chinese Pidgin English and Chinese lexicography on Southeast Asian languages. The second topic, however, deserves a separate volume.

Our focus on language has foregrounded the intimate contact between Southeast Asia's Chinese communities and the region's other ethnic groups, their languages, and their cultures. This topic has also proven popular in the field of Chinese-Malaysian (馬華 *Mahua*) literature. In addition, it speaks to the field of Southeast Asian linguistics, both in terms of the region's under-studied Sinitic languages and its archaic and/or substandard varieties of non-Sinitic languages, such as Malay and Tagalog. In our view, Southeast Asians seen as "indigenous" constitute an important part of the region's Sinophone landscapes. The local East Timorese who picked up Hakka from their Chinese employers (Chapter 2) have counterparts elsewhere in the region. Mandarin, too, is hardly the exclusive domain of Chinese-descended Southeast Asians, even though many rely on it as the only Sinitic language for which educational resources are available. Throughout the region, Mandarin is also spoken – with varying degrees of fluency – by indigenous Southeast Asians working in the tourism sector, the entertainment industry, elite private schools, and domestic service. Indeed, Southeast Asia's Sinophone is destined to live on in new incarnations, complete with new linguistic expressions and cultural manifestations.

The focus of this volume, however, has been on the region's quickly disappearing historical varieties.[1] The language practices it has highlighted – as well as many more that remain undocumented – are dying out under the weight of standardism, linguistic repression, and a lack of self-confidence within the communities. The contributors of this volume have responded in creative

---

1 Although the book's chapters have only tangentially touched upon religious customs, this is an additional line of evidence in which the Sinophone comes to the fore in all its diversity, as emerged from some of the unpublished presentations and discussions held during the workshop.

ways to the scarcity of primary sources on Southeast Asia's Sinitic varieties. In addition to fieldwork, their research has benefitted from online dictionaries, Wikipedia pages in different Sinitic varieties, unpublished MA and PhD theses, academic exercises, UN reports, newspaper articles, and websites, apps, and Facebook groups used by the communities, in a multitude of languages. At the same time, this volume has made clear how much still needs to be done. This makes the recuperation of Southeast Asia's marginalized Sinitic voices all the more urgent, as they provide key ingredients to reconceptualize the Sinophone in all its diversity. Between the minutiae of assorted language practices – many of which are only seen as peripherally "Chinese" – and more legitimized "texts" lies its true potential as a conceptually enriching framework.

Index

The letter *f* following a page number denotes a figure.