

**NANYANG  
TECHNOLOGICAL  
UNIVERSITY**  

---

**SINGAPORE**

**ASPECTS OF CAMBODIAN TEOCHEW  
GRAMMAR:  
A RADICAL CONSTRUCTION GRAMMAR  
ACCOUNT**

JOANNA ROSE MCFARLAND

SCHOOL OF HUMANITIES

2022

*ASPECTS OF CAMBODIAN TEOCHEW GRAMMAR:  
A RADICAL CONSTRUCTION GRAMMAR ACCOUNT*

Joanna Rose McFarland


School of Humanities

A thesis submitted to the Nanyang Technological University  
in partial fulfilment of the requirement for the degree of  
Doctor of Philosophy

*2022*



## Statement of Originality

I certify that all work submitted for this thesis is my original work. I declare that no other person's work has been used without due acknowledgement. Except where it is clearly stated that I have used some of this material elsewhere, this work has not been presented by me for assessment in any other institution or University. I certify that the data collected for this project are authentic and the investigations were conducted in accordance with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

Date:	October 26, 2022
Name of student:	Joanna Rose McFarland
Signature of student:	

## Supervisor Declaration Statement

I have reviewed the content of this thesis and to the best of my knowledge, it does not contain plagiarised materials. The presentation style is also consistent with what is expected of the degree awarded. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accordance with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

Date:	October 27, 2022
Name of Supervisor:	Alexander R. Coupe
Signature of Supervisor:	 30/10/2022
Name of Co-Supervisor:	Randy J. LaPolla
Signature of Co- Supervisor:	 2022-10-30


## Authorship Attribution Statement

Please select one of the following

This thesis contains material from 1 paper published in the following book in which I am listed as an author.

Portions of Chapters 2, 8, and 9 are published as McFarland, Joanna R. (2021).

Language contact and lexical changes in Khmer and Teochew in Cambodia and beyond. In T. Hoogervorst & C. Chia (Eds.), *Sinophone Southeast Asia* (pp. 91-128). Brill. [https://doi.org/10.1163/9789004473263\\_005](https://doi.org/10.1163/9789004473263_005)

Date:	October 26, 2022
Name of student:	Joanna Rose McFarland
Signature of student:	

## ACKNOWLEDGEMENTS

When I had my son, Ari, in April of 2021, some people seemed to express wonder that I was able to get a PhD and have a baby at the same time. At that point in time though, I hadn't yet *gotten* a PhD. I was merely in the process of doing it, which didn't seem that challenging. Actually *finishing* the dissertation that was required to get said PhD, well that indeed ended up being quite a feat. It would not have happened without the many people who have helped me along the way.

I could not have done this without the continued, unwavering support of my loving husband, Jesse. Not only was he willing to move our lives across the world for this venture, he also didn't hesitate to back me emotionally, morally, and financially through it all. I really can't thank him enough.

I greatly appreciate 公公 *kɔŋ<sup>55-11</sup>kɔŋ<sup>55</sup>* and 阿嬷 *a<sup>33</sup>ma<sup>52</sup>* for all their help, not limited to taking me around Cambodia, connecting me with several speakers and interpreting and translating data, answering my questions, providing data and recordings of their own, and more recently letting us move in with them, cooking us many delicious meals, and taking care of Ari.

Thank you to the rest of my family members, and each of our friends who supported, encouraged, and believed in me all this time. My little Ari was also a great motivator for me to finally get the work done and not push for another extension.

A heartfelt 致谢 *tɕɔi<sup>11</sup>sia<sup>11</sup>* to all my consultants and their families in Cambodia for graciously spending time with me, allowing me to record their language, and answering my questions. Special thanks to 老姨 *lau<sup>24-11</sup>i<sup>55</sup>* and 老丈 *lau<sup>24-11</sup>ti<sup>24</sup>* for hosting me (and Jesse) on several of my trips and providing many 好食 *hɔ<sup>52-24</sup>tɕia<sup>25</sup>* Cambodian/Teochew dishes and of course an abundance of មង្គុត *maŋ<sup>11</sup>k<sup>h</sup>u<sup>2</sup>*, សាវម៉ែវ *sau<sup>33</sup>mau<sup>33</sup>*, 涂莲 *t<sup>h</sup>ɔu<sup>11</sup>liɑŋ<sup>55</sup>*, and more.

Many thanks to my committee members: Randy LaPolla, for his depth of knowledge on Sinitic languages and functional linguistics that was clear in his always-prompt answers to my many

inquiries, and his insightful comments and suggestions on earlier drafts; Alec Coupe, for stepping up to be my main supervisor when required; Stephen Matthews, for sticking with me from my Master's days at HKU, and lending his expertise on Teochew; and Jingxia Lin, for jumping on my committee to support me when needed at the eleventh hour.

I am also indebted to Bill Croft for providing a pre-print draft of his forthcoming book (*Morphosyntax: Constructions of the World's Languages*), which Chapters 5, 6, and 7 of this dissertation are largely based on, and certainly would not have been the same without. Additional thanks are necessary for his extremely helpful comments on a draft version of this thesis.

I would also like to acknowledge the numerous people I met and befriended in Singapore along the way who made the country an unforgettable home for four years including: all the members of NTU Aquathlon, locals and exchange students alike, with whom I exhausted countless hours with (#IronProject2049 anyone?); Carmen and Lu Lin, my fellow new mommy linguists; the SUAW crew, Yvonne, Liliana, Aimee, Katherine, Joy, Christine, Ritu, and others, with whom I spent many hours of writing and more; my fellow PHP MCs of 2020-21 and each of the volunteers; and all of our other friends on the island, including Jesse's coworkers, who welcomed, supported, and entertained us over the years.

Thanks to Karen Zagona from my undergrad days at the University of Washington. While this dissertation certainly differs from the Chomskyian linguistics I worked on back then, I would not have pursued linguistics at a higher level without her support and encouragement.

Final shout out to Spirit and Captain who were always there for comfort and (usually forced) snuggles.

# TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS</b> .....	<b>i</b>
<b>TABLE OF CONTENTS</b> .....	<b>iii</b>
<b>LIST OF ABBREVIATIONS</b> .....	<b>vi</b>
<b>LIST OF TABLES</b> .....	<b>vii</b>
<b>LIST OF FIGURES</b> .....	<b>ix</b>
<b>SUMMARY</b> .....	<b>x</b>
<b>Chapter 1 – Introduction</b> .....	<b>1</b>
<b>1.1 Background and research questions</b> .....	<b>1</b>
<b>1.2 Radical Construction Grammar (RCG) framework</b> .....	<b>2</b>
1.2.1 Constructions are the basic units of grammar.....	3
1.2.2 Constructions are a pairing of form and function .....	9
1.2.3 There are no universal constructions .....	15
1.2.3 Summary.....	16
<b>1.3 Teochew language</b> .....	<b>16</b>
1.3.1 Genetic affiliation .....	16
1.3.2 Typological profile of Cambodian Teochew .....	19
<b>1.4 Fieldwork and data collection</b> .....	<b>21</b>
1.4.1 Primary data sources .....	21
1.4.1.1 Participant recruitment.....	22
1.4.1.2 Data collection procedures .....	24
1.4.2 Secondary data sources.....	30
<b>1.5 Data presentation</b> .....	<b>32</b>
<b>1.6 Organization of the thesis</b> .....	<b>35</b>
<b>Chapter 2 – Ethnic Chinese in Cambodia and Beyond</b> .....	<b>37</b>
<b>2.1 Terminology used for the Cambodian Chinese</b> .....	<b>37</b>
<b>2.2 History of the Chinese in Cambodia</b> .....	<b>39</b>
<b>2.3 Current Chinese resurgence</b> .....	<b>41</b>
<b>2.4 Sociolinguistic situation of the Teochew in Cambodia</b> .....	<b>43</b>
<b>2.5 Sociolinguistic situation of the Teochew in Southeast Asia</b> .....	<b>46</b>
<b>2.6 Chinese/Teochew influence on languages of Southeast Asia</b> .....	<b>49</b>
<b>2.7 Summary</b> .....	<b>53</b>
<b>Chapter 3 – Prior Teochew Language Research</b> .....	<b>54</b>
<b>3.1 Studies on Chaoshan varieties</b> .....	<b>54</b>
3.1.1 Missionary sources .....	54
3.1.2 Recent phonetic/phonological studies .....	55
3.1.3 Recent grammatical studies .....	58
<b>3.2 Studies on Southeast Asian varieties</b> .....	<b>60</b>
3.2.1 Cambodia.....	60
3.2.2 Indonesia .....	61

3.2.3 Malaysia.....	61
3.2.4 Singapore.....	65
3.2.5 Thailand.....	66
<b>3.3 Summary.....</b>	<b>67</b>
<b>Chapter 4 – Phonology.....</b>	<b>68</b>
<b>4.1 Consonants.....</b>	<b>68</b>
<b>4.2 Vowels.....</b>	<b>74</b>
4.2.1 Monophthongs.....	74
4.2.1 Diphthongs.....	78
4.2.3 Triphthongs.....	79
<b>4.3 Tone.....</b>	<b>79</b>
4.3.1 Citation tone.....	79
4.3.2 Tone sandhi.....	84
<b>4.4 Syllable Structure.....</b>	<b>88</b>
<b>4.5 Summary.....</b>	<b>91</b>
<b>Chapter 5 – Reference and Referent Expressions.....</b>	<b>95</b>
<b>5.1 CT Object Reference Constructions.....</b>	<b>96</b>
5.1.1 CT Common ObjR Constructions.....	96
5.1.2 CT Proper ObjR Constructions.....	97
5.1.3 CT Contextual ObjR Constructions.....	99
5.1.3.1 CT Personal Contextual ObjR Constructions.....	99
5.1.3.2 CT Deictic Contextual ObjR Constructions.....	109
<b>5.2 Information Status Continuum.....</b>	<b>114</b>
5.2.1 CT Identity Known Constructions.....	117
5.2.1.1 Active (in focus).....	119
5.2.1.2 Semi-active (accessible).....	120
5.2.1.3 Inactive and Inferable.....	126
5.2.2 CT Identity Unknown Constructions.....	129
5.2.3 CT Type Identifiable Constructions.....	134
5.2.4 CT Generic ObjR Constructions.....	137
<b>5.3 Extended Animacy Hierarchy.....</b>	<b>141</b>
<b>5.4 CT Property Reference Constructions.....</b>	<b>143</b>
<b>5.5 CT Action Reference Constructions.....</b>	<b>144</b>
<b>5.6 Summary.....</b>	<b>155</b>
<b>Chapter 6 – Modifiers and (Referent) Modification Constructions.....</b>	<b>156</b>
<b>6.1 CT Property Modification Constructions.....</b>	<b>158</b>
<b>6.2 Selecting modifiers.....</b>	<b>168</b>
6.2.1 CT Enumeration Modification Constructions.....	170
6.2.2 CT Quantifier Constructions.....	192
<b>6.3 CT Object Modification Constructions.....</b>	<b>198</b>
6.3.1 CT Anchoring Constructions (situating).....	198
6.3.1.1 CT Possessive Modification Constructions.....	199
6.3.1.2 CT Locative Modification Constructions.....	209
6.3.2 CT Typifying Constructions (subcategorizing).....	211
<b>6.4 CT Action Modification Constructions.....</b>	<b>215</b>
<b>6.5 Summary.....</b>	<b>222</b>

<b>Chapter 7 – Predication and Clauses .....</b>	<b>224</b>
<b>7.1 CT Action Predication Constructions .....</b>	<b>224</b>
7.1.1 CT Declarative Negation Constructions .....	233
7.1.1.1 唔 m <sup>11</sup> ‘NEG’ .....	233
7.1.1.2 无 bo <sup>55</sup> ‘NEG’ .....	237
7.1.1.3 未 bue <sup>11</sup> ‘not yet’ .....	239
7.1.1.4 𠵿 boi <sup>24</sup> ‘NEG.can’ .....	241
7.1.2 CT Interrogative Constructions .....	241
7.1.3 CT Imperative-Hortative Constructions .....	247
7.1.4 CT Complex Predicate Constructions .....	250
7.1.4.1 CT Eventive Complex Predicate Constructions .....	251
7.1.4.2 CT Stative Complex Predicate Constructions .....	262
7.1.5 Complex Sentences .....	264
7.1.5.1 CT Coordinate Clause Constructions (and CT Coordinate Constructions) .....	265
7.1.5.2 CT Equative and Similative Constructions .....	274
7.1.5.3 CT Conditional Constructions .....	275
<b>7.2 CT Object Predication Constructions .....</b>	<b>277</b>
7.2.1 CT Predicational Constructions .....	277
7.2.2 CT Nonpredicational Clause Constructions .....	281
<b>7.3 CT Property Predication Constructions .....</b>	<b>286</b>
<b>7.4 CT Locative Predication Constructions .....</b>	<b>292</b>
<b>7.5 Summary .....</b>	<b>296</b>
<b>Chapter 8 – Khmer Loanwords .....</b>	<b>297</b>
<b>8.1 Methodology .....</b>	<b>297</b>
<b>8.2 Loanwords .....</b>	<b>298</b>
<b>8.3 Integration of Loanwords .....</b>	<b>301</b>
<b>8.4 Speaker Variation .....</b>	<b>305</b>
<b>8.5 Discussion .....</b>	<b>307</b>
<b>Chapter 9 – Conclusion .....</b>	<b>315</b>
<b>9.1 Summary .....</b>	<b>315</b>
<b>9.2 Limitations and future work .....</b>	<b>318</b>
<b>9.3 Implications .....</b>	<b>326</b>
<b>References .....</b>	<b>328</b>

## LIST OF ABBREVIATIONS

1	first person	MOD	modifier marker
2	second person singular	N	native
3	third person singular	NEG	negative marker
A	agent/agent-like participant of transitive predicate	NMLZ	nominalizer
ATP	action-taking predicate	NOM	nominative
B	basic	ObjR	object reference
CLF	classifier	P	patient/patient-like participant of transitive predicate
COMP	complementizer	PASS	passive
COMP.DIR	directional complement	PDCO	Proximal Deictic Contextual ObjR
COP	copula	PFV	perfective aspect
CT	Cambodian Teochew	PL	plural
DDCO	Distal Deictic Contextual ObjR	POSS	possessive marker
E	elicitation	POT	complement of potential marker
EC	Enumeration Construction	Pr	proficient
EXCL	exclusive	PROG	progressive
EXLA	exclamative	PropM	property modification
F	fluent	PRT	particle
F#	female	PS	picture stimuli
G#	generation	RCG	Radical Construction Grammar
GEN	genitive	RED	reduplication
Hz	Hertz	REL	relativizer
I	intermediate	RC	resultative complement
IK	identity known	S	single participant role of intransitive predicate
INT	intensifier	SEQ	sequentiality marker
INTJ	interjection	SG	singular
IPA	International Phonetic Alphabet	SIM	simultaneity marker
JBT	Johor Bahru Teochew	ST	Singapore Teochew
LOC	locative		
LW	linker word		
M#	male		

## LIST OF TABLES

Table 1.1: English examples of information packaging of different semantic classes (Croft, forthcoming, p. 12).....	10
Table 1.2: Grammatical constructions combining semantic class (meaning) and propositional act functions (information packaging) .....	12
Table 1.3: RCG analysis of She is intelligent, an example of the English Property Predication Construction (Croft, forthcoming, p. 24) .....	15
Table 1.4: Primary consultant profiles .....	22
Table 1.5: Responses from Facebook participant recruitment .....	24
Table 1.6: Inventory of primary recording sources .....	29
Table 1.7: Secondary consultant profiles and recording list .....	30
Table 2.1: Non-student survey: Home language use (Bourgerie, 2017, p. 173).....	45
Table 2.2: Student survey: Home language use (Bourgerie, 2017, p. 173) .....	45
Table 2.3: Teochew appellation loanwords in Khmer and Thai .....	52
Table 4.1: Consonant phonemes in Cambodian Teochew.....	69
Table 4.2: Vowel phonemes in Cambodian Teochew .....	74
Table 4.3: Diphthongs in Cambodian Teochew.....	78
Table 4.4: Triphthongs in Cambodian Teochew.....	79
Table 4.5: The eight basic tones of Cambodian Teochew.....	80
Table 4.6: Tone sandhi in Cambodian Teochew .....	85
Table 4.7: Syllable structure examples in Cambodian Teochew.....	90
Table 5.1: Semantics of reference to an individual (adapted from Croft, forthcoming, p. 59) ..	96
Table 5.2: CT Personal Contextual ObjR Constructions.....	99
Table 5.3: Types of CT Contextual ObjRs (adapted from Croft, forthcoming, p. 63).....	113
Table 5.4: Information statuses and associated constructions (adapted from Croft, forthcoming, p. 66).....	115
Table 5.5: Accessibility Scale for referring expressions in Cambodian Teochew (adapted from Croft, forthcoming, p. 71) .....	118
Table 6.1: CT Property Modification Constructions, adapted from Croft (forthcoming, p. 95) .....	158
Table 6.2: CT Constructions of the selecting subfunction of the modification function, adapted from Croft (2021, p. 95).....	169
Table 6.3: Some CT Cardinal Enumeration Constructions .....	171
Table 6.4: Sortal classifiers in Cambodian Teochew .....	173

Table 6.5: CT Object Modification Constructions of the situating subfunction of the modification function, adapted from Croft (forthcoming, p. 95) .....	199
Table 6.6: Singapore Teochew Possessive Contextual Object Modifiers used for kinship terms (Low, 2014, p. 19) .....	209
Table 7.1: CT Interrogative Contextual ObjRs.....	244
Table 8.1 Khmer Loanwords in Cambodian Teochew .....	301
Table 8.2 Breakdown of the count of speakers using each word.....	307

## LIST OF FIGURES

Figure 1.1: Representation of action predication categories in Radical Construction Grammar (adapted from Croft, 2005, p. 285) .....	7
Figure 1.2: Venn diagram representing grammatical generalizations of English action predication behavior (Croft, 2020a, p. 187).....	8
Figure 1.3: The structure of a construction (Croft, forthcoming, p. 6).....	9
Figure 1.4: Sub-grouping of 12 Southern Min dialects (Kwok, 2018, p. 157).....	18
Figure 1.5: Simplified tree of Teochew’s genetic affiliation.....	19
Figure 1.6: Example text of a recruiting direct message.....	23
Figure 1.7: Family problem (San Roque et al., 2012) picture order arranged by G3F2 and G3F3 .....	27
Figure 4.1: Spectrograms comparing CT speaker’s (G1F4) 多 [tɕoi <sup>11</sup> ] ‘many’ to a ST speaker’s 多 [tsɔi <sup>11</sup> ] ‘many’ .....	71
Figure 4.2: Spectrogram of G1F4 热 /dzua <sup>25</sup> / ‘hot’ .....	72
Figure 4.3: Graphs showing vowel spaces (F <sub>1</sub> x F <sub>2</sub> ) for eight primary speakers.....	76
Figure 4.4: Graphs showing citation tone pitch (Hz) x time (seconds) for eight primary speakers .....	82
Figure 4.5: Pitch (Hz) x duration (ms) graph of the yin qu tone from twelve speakers in Hong’s (2013, p. 241) study.....	83
Figure 4.6: Graphs showing citation and sandhi tone comparisons for eight speakers .....	88
Figure 4.7: Syllable structure in Cambodian Teochew.....	89
Figure 5.1: Extended Animacy Hierarchy .....	142
Figure 6.1: The classifier indexical strategy (Croft, forthcoming, p. 115) .....	172
Figure 6.2: The linker strategy for relating two concepts (Croft, forthcoming, p. 123).....	181
Figure 7.1: The Cat Story context for example (50a).....	258
Figure 7.2: The Frog Story context for example (50b) .....	258

## SUMMARY

Teochew people, who originated in Chaoshan, China, have resided in Cambodia for hundreds of years, but limited work has been done on their language. This study attempts to rectify that by describing and documenting the Cambodian Teochew language, using the framework of Radical Construction Grammar (Croft, 2001, 2005, 2013, 2020b, forthcoming), while also examining the sociolinguistic situation of the Teochew in Cambodia. Additionally, as Teochew people have settled in many countries around the world, work was done to compare Cambodian Teochew to the varieties spoken in Chaoshan and Southeast Asia, with special attention paid to divergences, which were analyzed with respect to their relation to Khmer grammar.

The lexicon shows an influence from Khmer in the use of lexical borrowings for foodstuffs as well as some functional morphemes such as 𑜀𑜢𑜤𑜰𑜫 *hai<sup>33</sup> /hei<sup>33</sup>* ‘and’ and 𑜀𑜢𑜤𑜰𑜫 *pi<sup>33</sup>* ‘from’. In the phonology, divergences were found in the use of alveolo-palatal affricates (rather than alveolar affricates), reduced vowel nasalization, and reduced tones in sandhi environments for younger speakers. The morphosyntax showed some word order differences, such as post-object modification in CT Enumeration Modification Constructions, CT Quantifier Constructions, and CT Action Modification Constructions. There was also some evidence for path words such as 顶畔 *teŋ<sup>52-11</sup>paŋ<sup>55</sup>* ‘atop’ appearing before the ground, an order found in Khmer and not in other Teochew varieties. While these findings would be strengthened with additional data, they still have implications for the discussions surrounding language change in contact environments.

## Chapter 1 – Introduction

### 1.1 Background and research questions

Spoken vernaculars of Chinese are often referred to as ‘dialects’. However, due to their mutual unintelligibility, Chinese varieties may more properly be called ‘Sinitic languages’. Even still, historically scholars have attributed language differences to phonology and vocabulary, some viewing the syntax across Chinese as more or less the same, as evidenced by Chao’s notion of “one universal Chinese grammar” (1968, p. 13). Yue-Hashimoto (1993), Matthews (1999), and Chappell (2016), among others, have since disputed that idea, recognizing that differences in structural properties amongst Chinese varieties can have typological significance, thus giving rise to studies and descriptions of the grammar of a variety of languages of China, including Cai (2014), W. Chen (2011), Xu (2007), Zhang (2015) and the works within Chappell (2016).

It is also important to study varieties of the diaspora, as Chinese people have formed vibrant communities throughout Southeast Asia and around the world in which their Chinese culture and language is still very much alive today. One example of this is the Teochew people, who originated from the Chaoshan (Chaozhou-Shantou) region of eastern Guangdong, China. Throughout the 18th to the 20th centuries, many Teochew people emigrated from China throughout Asia including to Cambodia, Hong Kong, Indonesia, Laos, Malaysia, Philippines, Singapore, Thailand, and Vietnam. Teochew (also called Chaozhou, Chiuchow, Swatow, or Teochiu) is a member of the Sinitic Southern Min dialect group (see §1.3.1) and varieties in and out of China have been featured in several studies over the years, which are explored in Chapter 3. However, the literature is still lacking overall. This is especially true for the Teochew of Cambodia where the group makes up an abundant majority of the 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 would expect Cambodian Teochew<sup>1</sup> to have some differences in its sound system (phonology), grammar (morphosyntax), and vocabulary (lexicon). This research aims to provide a comprehensive reference grammar of this language variety, while also examining the sociolinguistic situation of the Teochew in Cambodia (see Chapter 2). Furthermore, work will be done to compare Cambodian Teochew to the varieties spoken in Chaoshan and Southeast Asia. Divergences will be analyzed with respect to their relation to Khmer grammar with the purpose of determining what extent Khmer has influenced the grammar of Cambodian Teochew.

## 1.2 Radical Construction Grammar (RCG) framework

This reference grammar is written in the framework of Radical Construction Grammar (RCG) (Croft, 2001, 2005, 2013, 2020b, forthcoming<sup>2</sup>). The following sections provide an overview of the framework and the three assumptions of RCG:

- (i) Constructions are the basic units of grammar (§1.2.1);
- (ii) Constructions are a pairing of form (morphosyntactic structure) and function (meaning + information packaging) (§1.2.2);
- (iii) There are no universal constructions. Constructions are all language-specific (§1.2.3)

---

<sup>1</sup> 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 speakers in different regions of Cambodia, as Peng (2012) found in Indonesia with the Jambi and Pontianak varieties (see Chapter 3).

<sup>2</sup> Page numbers from Croft (forthcoming) are based on a draft from January 2021 and they may differ from the published version.

While I followed the framework for grammatical description as laid out in Croft (forthcoming), I diverged from him in regards to the terminology I used throughout this dissertation. Croft (forthcoming) deferred to traditional grammatical terminology for some terms including noun, verb, adjective, predicate nominal, predicate adjectival, etc. He does this to “reclaim the terms” for functional linguistics from structural linguistics (W. Croft, personal communication, June 5, 2021). However, this can be confusing for new readers of RCG. While Croft defines these terms early on in his book by saying things like nouns are “object words... in referring phrases” (forthcoming, p. 37), it can be nearly impossible for readers to disassociate from the ‘old’ definition of noun and its word class usage (more details on this distinction are given below). Therefore, I have tried to stick to functional labels for constructions and have changed Croft’s terminology accordingly.

#### 1.2.1 Constructions are the basic units of grammar

CONSTRUCTIONS are defined as conventional symbolic units (Langacker, 1987). They are any type of morphosyntactic structure, complex (like a sentence) or atomic (like a word). Grammatical structures/constructions are morphosyntactic because they involve both morphology, the analysis of the internal structure of words, and syntax, the analysis of the internal structure of sentences/utterances (Croft, forthcoming, p. 4). For example, to analyze the English Enumeration Modification Construction in (1), we must pay attention to its syntax (the order of the number and object reference) and morphology (the singularity or plurality of the object).

(1) **English Enumeration Modification Construction**

one tree  
two tree-s  
three tree-s  
etc.

The English Enumeration Modification Construction can be represented schematically as in [NUM OBJR-NMB] where NUM, OBJR, and -NMB stand for roles in the construction which can be either types of words (NUM for *one*, *two*, *three*, etc; OBJR for *tree*, etc.) or bound morphemes (-NMB for the singular/plural suffix) (adapted from Croft, forthcoming, p. 4). But these roles cannot be taken outside of the constructions they are in because in RCG, constructions are the basic units of grammar, not the positions within the construction or the functions associated with them.

Taking constructions as the primitive elements of morphosyntactic representation is contrary to the approach found in traditional grammar, structuralist linguistics, and generative linguistics, where instead word classes or parts of speech are taken as basic units in morphosyntactic analysis. In this word class approach, words are placed into categories like noun, verb, adjective, etc. and what word class a word belongs to is considered an inherent property of that word. Croft (2001, 2013, 2020b) also calls this the building block model of syntax. This model is a reductionist theory where the analyses start with the smallest units (the word classes or parts of speech) and build up. RCG is contrary to this and is a nonreductionist theory. Instead, grammatical analyses start big (with constructions) and then define the smaller parts based on how they relate to the larger units.

The RCG approach is superior for several reasons. First, the word class or building block approach falls apart when we look at the empirical evidence. Word classes are defined by

the distributional method: analyses will look at what constructions syntactic categories (word classes) appear in to define that category. But constructions are not consistent across languages. For example, in many analyses, word classes or parts of speech are defined by morphological inflections. A word is a noun because of case markings or a verb because of person indexation. But defining a noun or verb by morphological inflections doesn't hold for analytic languages such as Vietnamese and Chinese. For those languages, linguists have to look for other constructions that allow them to define a noun or a verb in another way. This does not look like there is a meaningful way to define the categories of nouns and verbs across languages.

The way the distributional method is applied in the building block model does not even work within languages. When looking at English, there is a contrast between what can appear in the English Active Voice Construction in (2) and the English Passive Voice Construction in (3).

(2) **English Active Voice Construction**

- a. John kicked **the ball**.
- b. John weighs **160 pounds**.

(3) **English Passive Voice Construction**

- a. **The ball** was kicked by John.
- b. **\*160 pounds** were weighed by John.

While the (a) examples show a correspondence between the bolded Active Direct Object category and the Passive Subject category, the (b) examples contradict this. Yet many linguists still analyze the Passive Subject as criterial for defining the Active Direct Object category, despite these distributional facts (i.e., they claim that *160 pounds* is not an Active Direct Object).

Arbitrarily selecting constructions to define the same word class across languages and ignoring distributional facts within one language are examples of what Croft (2001, 2013, 2020b) calls *METHODOLOGICAL OPPORTUNISM*. Linguists will ignore the empirical evidence in order to maintain the building block model of syntax and the idea that there are consistent categories such as noun and verb within and across languages. In reality, the sentences in (2) and (3) tell us more about what types of Subjects can occur in the English Passive Voice Construction rather than the category of Active Direct Object.

Also at issue with the word class approach is that it presents a circular argument (Croft, 2001, 2005, 2020a, 2020b). With the distributional method, a word class/category is defined by being part of certain constructions. For example, linguists might say that the word class of adjective is defined by being in the Object Modification Construction, as in (4).

- (4) **Object Modification Construction**
- a. a **blue** flower
  - b. the **big** dog

But then at the same time, the word classes/categories are taken as primitives or building blocks and constructions themselves are defined by being built out of these units. So the Object Modification Construction is defined as being made up of an article, an adjective, and a noun. This is circular. With this type of reasoning, all one is really saying is that ‘an adjective occurs in a construction used to define an adjective’. This doesn’t tell us anything about the category of ‘adjective’. Another example of this circular argument is that linguists will often define a noun phrase as a phrase headed by a noun, and at the same time being the head of a noun phrase is criterial for nounhood.

To avoid this circular argument, we either have to discard the distributional method, or we discard the idea that syntactic categories are primitives (Croft, 2005). There is nothing inherently wrong with the distributional method. At issue is linguists who selectively use the distributional method to define word classes in an ad hoc manner as described above. But distributional analysis shows us empirical facts of language. It doesn't make sense to discard the distributional method. Rather we should rid ourselves of the notion that syntactic categories are primitives. Instead, it is constructions that are primitives. Morphosyntactic categories only exist in the constructions in which they are defined. For example, there is the English Transitive Action Predicate in the English Transitive Construction, the English Intransitive Action Predicate in the English Intransitive Construction, etc. This is not to say that there is not a relation between the Action Predicates in each of these constructions, as we can't deny that they have some of the same behavior in terms of inflections. Croft (2005, pp. 284-5, 2013, p. 9) showed how this can be represented in RCG by positing a category that is defined by tense-agreement inflection. We'll call this MActPred (morphological action predicates). The taxonomic relationship between MActPred and Intransitive Action Predicates and Transitive Action Predicates in RCG is represented in Figure 1.1.

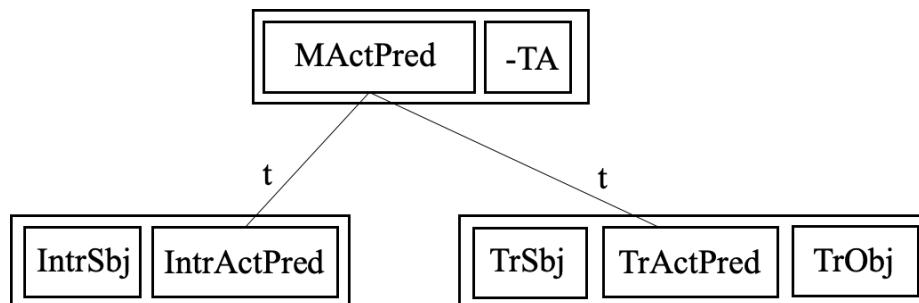


Figure 1.1: Representation of action predication categories in Radical Construction Grammar (adapted from Croft, 2005, p. 285)

Croft (2013, 2020a) demonstrated that alternative methods can be used to represent syntactic generalizations besides a taxonomy like Figure 1.1. For example, we could also represent the same concept spatially, like in the Venn diagram in Figure 1.2.

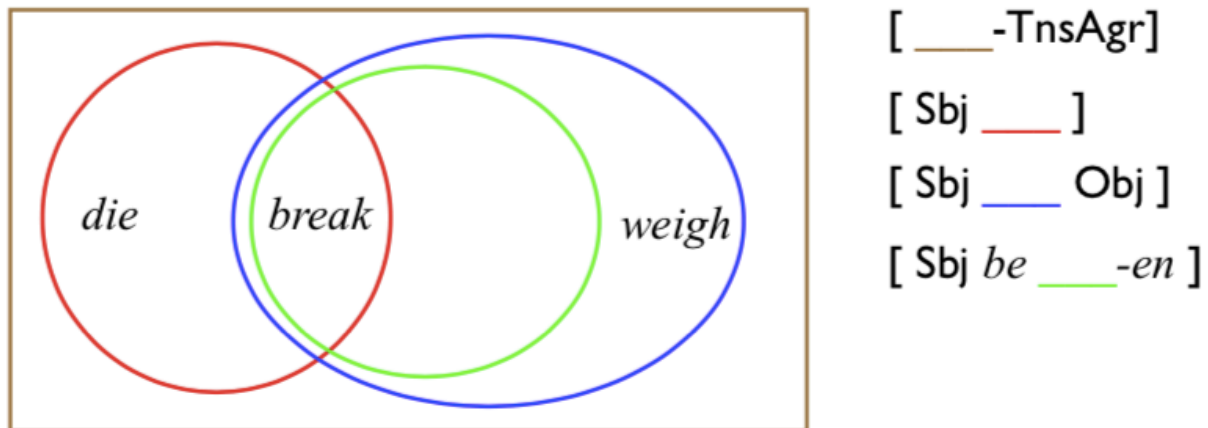


Figure 1.2: Venn diagram representing grammatical generalizations of English action predication behavior (Croft, 2020a, p. 187)

The brown box represents all English Morphological Action Predicates (MActPred) which are characterized by being inflected with tense and agreement. The red circle containing *die* and *break* represents English Intransitive Action Predicates used in the English Intransitive Construction. *Break* is also in the green circle which represents English Action Predicates that occur in the English Passive Construction. Finally, *break* and *weigh* are part of the blue circle which represents the English Transitive Action Predicates that are found in the English Transitive Construction.

No matter the way of looking at such grammatical generalizations within a language (and there are more than the two ways shown above), we are still analyzing and representing such generalizations in terms of constructions, rather than in terms of features and values

defined separately from the constructions. In RCG, constructions are always the basic units of grammar and a construction's individual elements cannot be defined outside of that construction.

### 1.2.2 Constructions are a pairing of form and function

Constructions must be analyzed with respect to how form expresses function, including both meaning and information packaging. The idea that constructions are made up of form (morphosyntax) and function (information structure and semantic content) is represented in Figure 1.3.

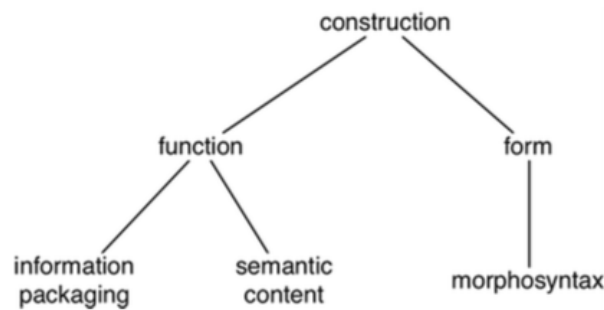


Figure 1.3: The structure of a construction (Croft, forthcoming, p. 6)

The separation of semantic content and information packaging is important when analyzing the function of a construction because in theory any semantic content can be put into any type of information packaging. Words can be classified by semantic labels such as objects (*tree, bush*), properties (*tall, green*), and actions (*fall, grow*). While in sentences like *The tall tree fell* or *A green bush grew*, the same words can be categorized as referring (*tree, bush*), modifying (*tall, green*), and predicating (*fell, grew*). The latter categorization is the information packaging of the semantic content, which can be elaborated as follows (Croft, forthcoming, p. 13):

**reference** – what the speaker is talking about

**modification** – further information on the referent

**predication** – what the speaker is declaring about the referents in a certain utterance

These three crucial information packaging functions are also called PROPOSITIONAL ACT FUNCTIONS (Croft, forthcoming, p. 13).

The ways the different semantic classes (objects, properties, and actions) can participate in information packaging/propositional act functions (reference, modification, and predication) are showcased with examples in Table 1.1.

	<b>reference</b>	<b>modification</b>	<b>predication</b>
<b>object</b>	<i>the sharp thorns</i>	<i>the bush's thorns</i>	<i>It is a thorn.</i>
<b>property</b>	<i>sharpness</i>	<i>the sharp thorns</i>	<i>Those thorns are sharp.</i>
<b>action</b>	<i>I said [that the thorns scratched me]. the [scratching of the thorns]</i>	<i>the thorns [that scratched me] the thorns [scratching me]</i>	<i>The sharp thorns scratched me.</i>

Table 1.1: English examples of information packaging of different semantic classes (Croft, forthcoming, p. 12)

Table 1.1 uncovers more flaws in the traditional analyses where ‘nouns’ are objects, ‘adjectives’ are properties, and ‘verbs’ are actions because this isn’t always the case. We can see that properties and actions can also be referred to (*sharpness*, *scratching*), while objects and actions can also be modifying (*bush’s*, *scratching*), and objects and properties can be predicating (*is a thorn*, *are sharp*). Another issue is that it is often unclear whether an author is using word class or part of speech terminology to refer to a word’s semantic content or information packaging. For example, the term ‘demonstrative adjective’ can be defined as a ‘deictic word [“demonstrative”] used as a modifier [“adjective”]’, like *that* in *that book*, while ‘predicate adjective’ is defined as a ‘property word [“adjective”] used in predication

[“predicate”], like *That tree is tall* (Croft, forthcoming, pp. 14-5). In these examples, the term ‘adjective’ is used to describe the information packaging (modification) and the semantic content (property), respectively. Readers of a reference grammar that uses such terminology will have to use deduction to figure out exactly which way the author is using a term. At the same time, word class or part of speech terminology may be used in other confusing ways such as “Sidaama numerals are adjectives” (Kawachi, 2007, p. 135) and Li and Thompson’s (1981) use of the term “adjectival verb” when describing Mandarin. RCG’s separation of semantic content and information packaging gives us better terminology to explain these phenomena. Rather than claiming “Sidaama numerals are adjectives”, we could say “the Sidaama Modification Construction is also used for enumeration”. “Adjectival verbs” in Mandarin would be better characterized by their appearance in Mandarin Property Predication Constructions.

Acknowledging the two dimensions of function (meaning and information packaging) allows us to meaningfully compare cross linguistically without relying on traditional definitions of word classes or parts of speech which have been shown to be inconsistent. Instead, we can compare different languages’ approaches to the predication of property concepts or object concepts used in modification. These non-language-specific approaches are examples of Haspelmath’s (2010) COMPARATIVE CONCEPTS, which are needed for cross linguistic comparison. At its most basic, a comparative concept is functional, or solely related to meaning. Croft’s (2016) HYBRID COMPARATIVE CONCEPTS go a step further to combine form and function. One type of hybrid comparative concept is a construction, which has already been defined as a pairing of form and function (see Figure 1.3). When naming/labeling constructions, Croft (forthcoming) proposed using a combination of the meaning and

information packaging, such as ‘property predication construction’. However, in some cases he deferred to traditional grammatical terminology that has already been used to describe how a language expresses certain functions. Thus, Croft (forthcoming) used terms such as ‘predicate nominal construction’ rather than ‘object predication construction’ and ‘relative clause construction’ rather than ‘action modification construction’. As previously mentioned, this dissertation uses the meaning and information packing labels rather than traditional grammatical terminology. Following typological convention, there is an orthographic distinction between comparative concepts, which appear in all lowercase, and language-specific categories, which are labeled with initial capitalization (Bybee, 1985, p. 141; Comrie, 1976, p. 10; Croft, 2001, p. 12; Dahl, 1985, p. 34; Haspelmath, 2010, p. 674). Thus, an object predication construction is a comparative concept while the English Object Predication Construction is a construction specific to English. Table 1.2 showcases labeling conventions adapted from Croft (forthcoming) for the combinations of the three basic semantic classes (object, property, and action) and three fundamental propositional acts (reference, modification, and predication). This terminology is followed throughout this dissertation.

<b>Semantic Class</b>	<b>Propositional act</b>		
	<i>reference</i>	<i>modification</i>	<i>predication</i>
<i>object</i>	object reference (objR)	object modification	object predication
<i>property</i>	property reference	property modification (propM)	property predication
<i>action</i>	action reference	action modification	action predication

Table 1.2: Grammatical constructions combining semantic class (meaning) and propositional act functions (information packaging)<sup>3</sup>

---

<sup>3</sup> See Croft (forthcoming, p. 29) for a similar table showcasing his terminology.

While constructions describe **what** a language is expressing, they do not tell us **how** that function is expressed. Croft’s second hybrid comparative concept is a STRATEGY, which is a construction that expresses a certain function (the what), along with a certain grammatical form that can be defined cross-linguistically (the how). Several different types of strategies are used in language. ENCODING STRATEGIES represent different ways the same function is conveyed in multiple languages or different ways of conveying the same function in one language (Croft, forthcoming, p. 20). Examples of encoding strategies include order of elements, such as how property modification can be expressed with a pre-object strategy (PropM-ObjR) or post-object strategy (ObjR-PropM), or the use of overt coding such as an inflecting copula strategy (*is*) which is used in the English Property Predication Construction and the English Object Predication Construction. Such encoding strategies can be used to compare constructions in different languages or within one language. For example, we can say that the inflecting copula strategy is used in English but not in Mandarin in property predication constructions, as in (5).

- (5) a. He is fat.  
 b. 他 胖  
 Tā pàng  
 3sg fat  
 ‘S/he is fat.’

Meanwhile, looking at just English, we can find several strategies for expressing object modification as in (6).

- (6) a. the regulations of the university
- b. the university's regulations
- c. university regulations

At the same time, a language might have multiple constructions that use the same strategy, such as in Spanish, which has two inflecting copula strategies, due to the language's two copulas, *ser* and *estar*. Croft (forthcoming, p. 20) proposed differentiating these two strategies using language-specific construction names like the *Ser* Copula Construction and the *Estar* Copula Construction. CO-EXPRESSION STRATEGIES describe categories that are relevant across multiple constructions, such as the alignment strategies found in ergative and accusative languages which look at categories across transitive and intransitive constructions (Croft, forthcoming, p. 20). RECRUITMENT STRATEGIES describe a situation where a construction that is used for a certain function is recruited to be used in a different function (Croft, forthcoming, p. 21). An example of this is seen in Russian where the same construction that is used for predication of location is used for the predication of possession. This is not an exhaustive list of all the types of strategies used in language. Exploring both constructions and strategies provides us with the comparative concepts needed to talk about grammatical constructions across languages.

To sum up this section, Table 1.3 provides an RCG analysis of the English sentence *She is intelligent*, which is an example of the English Property Predication Construction.

<b><i>Sentence</i></b>	<i>She</i>	<i>is</i>	<i>intelligent</i>
<b>Semantics</b>	object		property
<b>Information Packaging</b>	reference		predication
<b>Construction</b>	property predication construction		
<b>Strategy</b>	inflecting copula ( <i>is</i> )		

Table 1.3: RCG analysis of *She is intelligent*, an example of the English Property Predication Construction (Croft, forthcoming, p. 24)

In terms of semantic content, *she* is an object and *intelligent* is a property. For information packaging, the referent is *she* while *intelligent* is the predicate. This construction is called the English Property Predication Construction due to its expression of property predication. The term for the construction has initial capitalization because it is referring to an English-specific construction. At the same time, we can use the term property predication construction as a comparative concept to compare English's expression of this function to other languages. In addition to *she* and *intelligent*, the English Property Predication Construction also contains the word *is*, which is an instance of the inflecting copula strategy.

### 1.2.3 There are no universal constructions

In RCG (and arguably in general), there are no universal constructions. Constructions are all language-specific. There are no constructions such as passive or coordination that are the same in all languages. To prove this definitively, we would have to go through every construction thought to be universal and prove that it is not. Croft (2005) argued that the onus of proof should be on the person claiming the existence of any universal construction. Nonetheless, he provides examples of voice constructions in several languages which provide empirical evidence against a universal structural description of passive or inverse constructions (Croft, 2005, pp. 304-8). While we can look at constructions across languages in a comparative

manner, we should never enter into an analysis of a new language with assumptions about what constructions it contains. And if we find that a language does have a certain construction, such as a passive construction, it is still language-specific. We can compare it to a passive construction in another language, but we can't assume that it is identical. This approach was followed in researching and writing this grammar of Cambodian Teochew. Constructions were analyzed when they were found, rather than entering into the analysis with a need to find certain constructions. If a construction was not found in my collected data (see §1.4), then it is not included in this reference grammar. That is not to say that it does not exist in Cambodian Teochew, just that it was not found in my research.

### 1.2.3 Summary

Traditional approaches to grammar writing often have linguists focused on word classes and parts of speech. In fact, grammars are usually arranged around such terminology, perhaps with entire chapters dedicated to nouns, verbs, adjectives, etc. in a language. This section aimed to showcase why this approach to language analysis is flawed and why an approach such as RCG is superior. This is why RCG was selected as a framework for this reference grammar. In terms of how RCG is applied, §1.6 covers the organization of the dissertation, including the types of constructions that will be explored in detail.

## 1.3 Teochew language

### 1.3.1 Genetic affiliation

Determining the genetic affiliation of Teochew and other Sinitic languages is not a simple task, though many scholars have tried. According to the Sino-Tibetan hypothesis, Sinitic and Tibeto-Burman share a common ancestor (Matisoff, 2003; Thurgood & LaPolla, 2003). Sinitic languages are then often sorted into seven major groups per Yuan (1983): Mandarin,

Wu, Xiang, Gan, Min, Yue, and Hakka (Keijia).<sup>4</sup> Norman (1988) grouped these seven into three supergroups of Northern (Mandarin), Central (Wu, Xiang, Gan), and Southern (Min, Yue, Hakka), but not all scholars take this approach. Dong (2014) set aside the question of a central and southern divide, instead putting Mandarin in a Northern group and the remaining six in a Southern branch. Determining a clear family tree structure in the traditional sense (Indo-European) is not an easy task in the Sinitic language family due to a history of frequent migration and language contact (LaPolla, 2010).

There is much diversity within the Min branch itself. Norman (1988) divided the group into two due to different historical migrations in the Min area: Western Min, made up of Far Western Min (Shaowu, Jiangle), Northwestern Min (Jianou, Jianyang, Chongan), and Central Min (Yongan); and Eastern Min, consisting of Southern Min (Xiamen, Chaozhou, Taiwanese, Hainan) and Northeastern Min (Fuan, Ningde, Fuzhou). Scholars such as Kwok (2018) have further split Southern Min as in Figure 1.4.

---

<sup>4</sup> This is not without exception as Thurgood (2003) subsumes Hakka under Gan.

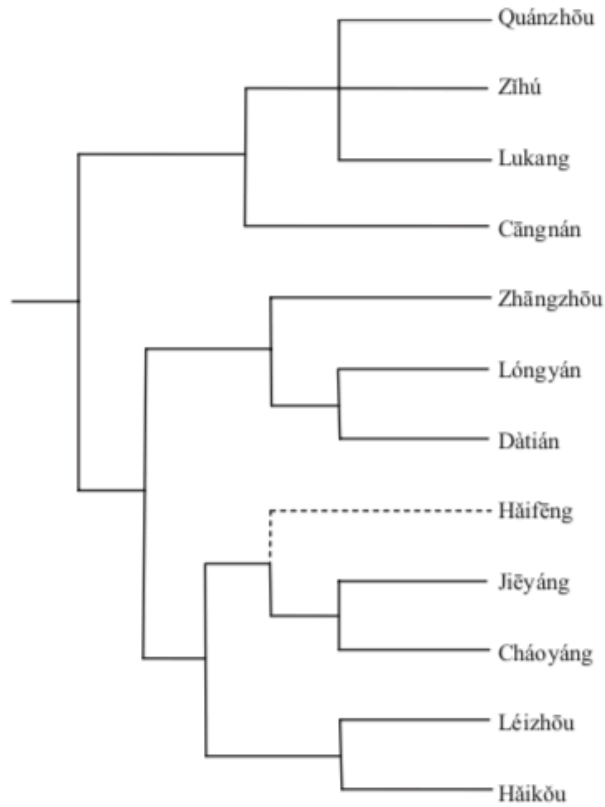


Figure 1.4: Sub-grouping of 12 Southern Min dialects (Kwok, 2018, p. 157)

Keeping the above information in mind and setting aside some of the complexities and controversies, Figure 1.5 showcases a simplified tree of Teochew’s hypothesized genetic affiliations.

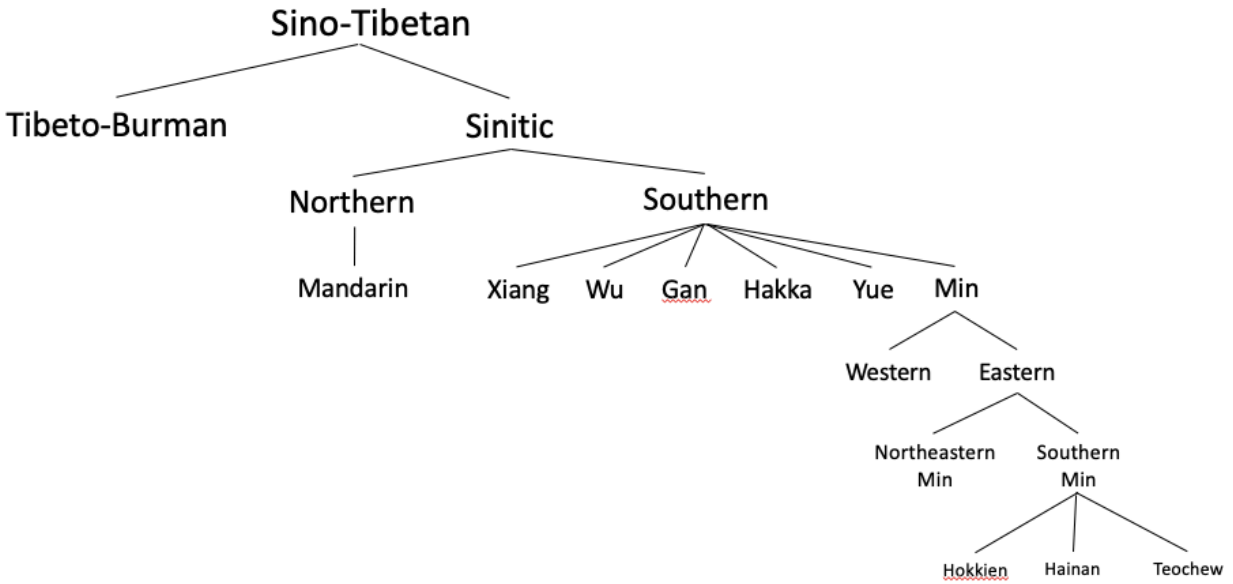


Figure 1.5: Simplified tree of Teochew’s genetic affiliation

### 1.3.2 Typological profile of Cambodian Teochew

The Cambodian Teochew phonology has a three-way contrast in velar and bilabial stops, but only two contrastive alveolar stops, /t/ and /t<sup>h</sup>/, which is consistent with other Teochew varieties (§4.1). There are also three alveolo-palatal affricates, as compared to other Teochew varieties’ alveolar affricates. The language does not contain consonant clusters and the only consonants allowed in the coda are /p/, /ʔ/, /k/, /m/ and /ŋ/ (§4.4). In words with historic coda /n/, nasalization has been absorbed into the vowel. This vowel nasalization is contrastive, though Cambodian Teochew seems to be showing reduced vowel nasalization as compared to other Teochew varieties (§4.2). The language has eight basic tones and makes productive use of tone sandhi whereby a syllable’s tone often differs in isolation versus in context (§4.3).

Like other Sinitic languages, the Cambodian Teochew morphology is characterized by limited affixation and frequent use of compounding and reduplication. Plural distinctions are

only found on CT Personal Contextual ObjR Constructions (§5.1.3.1). The inclusive and exclusive distinction for First Person Plural Contextual ObjR Constructions found in other Teochew varieties is seemingly reduced or altogether lost for some Cambodian Teochew speakers. Case and gender are not marked on CT Common ObjR or Personal Contextual ObjR Constructions. While special genitive forms used for kinship possession have been found in other Teochew varieties, none were found in the data set on Cambodian Teochew (§6.3.1.1).

Teochew varieties typically display pre-object modification but there is evidence of an increasing use of post-object modification in Cambodian Teochew which could potentially be explained by contact with Khmer, a language with frequent post-object modification. This type of modification is seen in CT Enumeration Modification Constructions (§6.2.1), CT Quantifier Constructions (§6.2.2), and CT Action Modification Constructions (§6.4).

Cambodian Teochew makes frequent use of sortal classifiers for indexing objects, such as in CT Enumeration Modification Constructions, though there is evidence that the classifier is no longer strictly required, as it typically is in Sinitic languages (§6.2.1). The classifier system also appears greatly simplified as compared to the Jieyang variety, and the use of the generic classifier 个 *ka<sup>55</sup>*, which we call a linker word, is increasingly prevalent. Classifiers, especially 个 *ka<sup>55</sup>*, are also used as for information status distinctions (§5.2).

There is no evidence for a syntactic pivot in Cambodian Teochew (§7.1). Word order is topic comment. Arguments are frequently elided if they can be recovered from the discourse context (§5.2.1). Some lexical predicates have grammaticalized into other functions, such as 𠵼 *hɔu<sup>11</sup>* ‘give’ which is seen in CT Causative Constructions and benefactive function (§7.1), and 𠵼 *tã<sup>21</sup>* ‘say’ which is used as a complementizer (§5.5). Cambodian Teochew makes frequent use of the serial predicate strategy for CT Complex Predicate Constructions, including those

showcasing motion events (§7.1.4.1) and depictive and manner functions (§7.1.4.2). Object predication uses a copula strategy (§7.2), while property predication does not, and instead resembles action predication (§7.3).

Given the language's lack of official status in Cambodia and its general relegation to the home domain, some speakers are showing signs of language attrition and deficient lexicons. As a result, lexical innovations in the form of compounded or juxtaposed words are common (§6.3.2; §6.4). Alternatively, Cambodian Teochew speakers code-switch, or borrow some words from Khmer, Mandarin, and English, though the language is generally resistive to lexical borrowings as compared to other Teochew varieties (see McFarland, 2021).

#### 1.4 Fieldwork and data collection

A total of eleven weeks of fieldwork were completed over five trips made to Phnom Penh in 2018-2020. The data gathered from nine speakers on these trips constitutes the bulk of this dissertation, described in (§1.4.1). However, upon analysis of this data, many holes were found, especially with regards to more rare constructions, limiting my ability to write a full and detailed grammar. Plans to return to Cambodia for further data collection in the spring and summer of 2020 were foiled by travel restrictions caused by COVID-19. Therefore, additional secondary data, described in (§1.4.2), has been pulled into this dissertation in order to make the grammar as complete as possible. The following sections outline the processes used for gathering the primary and secondary data.

##### 1.4.1 Primary data sources

Table 1.4 shows the breakdown of the nine primary consultants including a unique speaker ID as well as their age at the time of recording, generation (as defined in §1.4.1.2), gender, and language repertoire. Speaker ID is made up of the speaker's generation, gender,

and an assigned number. Detailed descriptions on speaker background and recruitment, as well as procedures follow the table.

Speaker ID	Age	Generation	Gender	Languages
G3F1	21	3	F	Khmer (N), English (F), Mandarin (F)
G3F2	16	3	F	Khmer (N), English (Pr), Mandarin (Pr)
G3F3	18	3	F	Khmer (N), English (Pr), Mandarin (Pr)
G2M1	55	2	M	Khmer (N)
G1M1	70	1	M	Khmer (N), Mandarin (F), French (I), Vietnamese (I), Thai (I), English (B)
G1F1	68	1	F	Khmer (N), Mandarin (Pr)
G1F2	65	1	F	Khmer (N)
G1F3	80	1	F	Khmer (N)
G1F4	65+	1	F	Khmer (N)

Table 1.4: Primary consultant profiles

#### 1.4.1.1 Participant recruitment

Five primary consultants (G2M1, G1M1, G1F1, G1F2, G1F4) were found through the researcher’s familial connections in Cambodia. G1M1 and G1F1 are a married couple living in the same household as their helper, G1F2, while G2M1 is the couple’s nephew, and G1F4 is a distant in-law relation (G2M1’s son’s grandmother-in-law). Thus while these five speakers are all in some way connected, I find their relations and upbringings are distant enough to bring a diversity of data.

The four remaining primary consultants (G3F1, G3F2, G3F3, and G1F3) were recruited via direct message requests targeted at Phnom Penh residents who were participants of Teochew-related Facebook pages including the group ‘Gaginang’ and the Cambodian Teochew

Association's (柬埔寨潮州会馆) business page.<sup>5</sup> Potential recruits were sent messages in English or in English and Mandarin asking for their participation. An example message is shown in Figure 1.6.

Hello! I am an American PhD student in Linguistics in Singapore. My husband is Cambodian Teochew American. I am looking for speakers of Teochew in Cambodia. Do you speak Teochew? Could you help me with my research? 你好！我是美国人，是博士学生，在新加坡学语言学。我的爱人是柬埔寨-潮州-美国人。我找在柬埔寨说潮州话的人。你说潮州话？你可以帮我吗？

Figure 1.6: Example text of a recruiting direct message

Messages were sent to nearly 50 people and ten responses were received (20%). Of those who responded, attempts were made to have brief chats with the respondents about their language background and if they knew anyone who spoke Teochew and could make an introduction. This information provided some additional insights into the sociolinguistic situation of the Cambodian Teochew which is explored in detail in Chapter 2. Results of the Facebook participant recruitment are found in Table 1.5.

---

<sup>5</sup> This page is offline at the time of writing (November 20, 2020).

Teochew speaker?	Do you know anyone who speaks Teochew?	Consultant connection	Other information
Yes	Family and friends	G3F3, introduction to G3F2	
No	Grandmother	Introduction to G1F3	Mandarin speaker
Yes	Friend and grandmother	Introduction to G3F1, Garden tour video (see §1.4.2)	
Yes	Unknown	N/A	Mandarin speaker
No	No	N/A	Mandarin speaker
No	Grandmother and friend	N/A	Mandarin speaker Attended Duan Hua Chinese school Identifies as Cambodian Teochew
No	Brother and sister	N/A	
No	Wife and relatives	N/A	
No	No	N/A	
Yes	Unknown	N/A	

Table 1.5: Responses from Facebook participant recruitment

One respondent was primary consultant G3F3 and she brought her friend G3F2 along to our scheduled interview. One respondent provided an introduction to his grandmother, speaker G1F3. Another respondent introduced me to speaker G3F1, and also provided me with the ‘Garden tour video’ which is used as a secondary data source (see §1.4.2). These speakers (G3F1, G3F2, G3F3, and G1F3) represent four unique families.

*1.4.1.2 Data collection procedures*

Informal interviews were conducted with each consultant, and any present family members, and 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 65 and older who were born in Cambodia and grew up before the wartime and Vietnamese occupation (1970-1989). 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 35-64. These people were children during the wartime. One male speaker was 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 have spoken Teochew and Khmer from a young age, while the three G3 speakers also speak English, and five consultants (G1M1, G1F1, G3F1, G3F2, G3F3) are proficient or fluent in Mandarin. Language proficiency in Table 1.4 is defined as: Native (N), Fluent (F), Proficient (Pr), Intermediate (I), Basic (B). Interviews, consent taking, and procedural instructions were conducted in English with G3 speakers and in a mixture of Khmer and Teochew with the G1 and G2 speakers 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. Ideally recording sessions would include video and audio in order to see some mouth and tongue movements, to more easily follow along with any visual stimuli, and to include any extra-linguistic information such as gestures and facial expressions. Eight of the nine primary consultants consented to video while speaker G3F1 only consented to audio recording.<sup>6</sup>

---

<sup>6</sup> Consent forms, following IRB Reference number: IRB-2020-01-020-01, were provided in English and Khmer, translated by Lingua Technologies International.

Data gathered is primarily staged communicative events (SCE) per Himmelmann (1998) with some elicitation. Limited amounts of observed communicative events (OCE) occur in some recordings from participant interactions and in speaker/interpreter interactions. Two early interviews (G1F1, G1M1) included elicitation using Swadesh’s (1955) list of non-cultural vocabulary. Words were provided in Mandarin and consultants were asked to translate them into Teochew. Some additional elicited data was collected from the three G3 speakers over Facebook Messenger. These speakers were sent words or phrases in English and they returned a voice recording in Teochew. In later interviews (all speakers except G1M1), consultants were shown picture stimuli of items and actions and asked to produce (an) appropriate word(s) or phrase(s) in Teochew. Image prompts consisted of local foods and dishes, items from Swadesh’s (1955) list, pictures representing words identified as potential minimal pairs, and selected items that were listed as loanwords in Singapore Teochew by Goh (2017) and Low (2014). The number of prompts ranged from 133 for earlier interviews to 233 for later interviews. Speakers were generally unprompted throughout the presentation of the picture stimuli, though occasional prompts in Teochew were provided by the researcher such as *tɕi<sup>51</sup>kaɪ<sup>55-11</sup> mi<sup>25</sup>kaɪ<sup>55</sup>* ‘what is this?’, *tɕi<sup>51</sup>kaɪ<sup>55-11</sup> mi<sup>25</sup>kaɪ<sup>55</sup> sek<sup>2</sup>* ‘what color is this?’, and *kuɪ<sup>51</sup>* ‘how many?’. Naturalistic texts were also recorded by having consultants describe scenes from picture books including *The Cat Story* by Katie Gao (N=3), *Fann* by Patrick Yee (N=1), *Frog, Where Are You?* by Mercer Mayer (N=1), *Looking for Felix* by Libby Hathorn (N=3), *The Red Book* by Barbara Lehman (N=1), *Zoom* by Istvan Banyai (N=4), and *While You Are Sleeping* by Mariana Ruiz Johnson (N=2). Participants were minimally instructed in either English by the researcher or in Khmer or Teochew by familial interpreters, to look at the

---

pictures and describe the scenes and stories using Teochew. Finally, consultants G3F2 and G3F3, who were interviewed together, performed the family problem picture task from San Roque et al. (2012). A set of sixteen pictures were individually presented to the speakers in San Roque et al.'s (2012) presentation order (pp. 142-3). Consultants were instructed in English to describe what they were seeing in each of the pictures. Next, they were given the full set of pictures and asked to work together and converse in Teochew to arrange them into an order of their choosing which would create a logical story. Lastly, they were told to tell the story as they had arranged it. Their order is presented in Figure 1.7.



Figure 1.7: Family problem (San Roque et al., 2012) picture order arranged by G3F2 and G3F3

Table 1.6 shows the inventory of recordings collected from primary consultants. In total, the primary data sources consist of approximately two hours of texts from picture books/stories (N = 16), two hours of data from picture stimuli, and thirty minutes of elicited data, including that from Facebook Messenger.

Speaker ID	Recordings		
	Name	Time (MM:SS)	ID
G3F1 [audio only]	Picture stimuli [N = 226]	12:23	PS
	Zoom	03:47	1
	Looking for Felix	01:50	2
	Cat Story	02:04	3
	Facebook Messenger		E
G3F2	Picture stimuli [N = 233]	25:15	PS
	Family problem	23:06	4
	Facebook Messenger		E
G3F3	Picture stimuli [N = 233]	25:15	PS
	Family problem	23:06	4
	Facebook Messenger		E
G2M1	Picture stimuli [N = 133]	07:11	PS
	Zoom	05:10	5
	Looking for Felix	02:29	6
G1M1	Swadesh elicitation	16:07	E
	Fann	29:52	7
G1F1	Swadesh elicitation	09:04	E
	Picture stimuli [N = 133]	08:05	PS
	Zoom	10:03	8
	Cat Story	03:10	9
	While You Are Sleeping	07:57	10
	Frog	05:57	11
	Red Book	05:00	12
G1F2	Picture stimuli [N = 133]	15:24	PS
	Looking for Felix	09:09	13
	Cat Story	04:16	14
G1F3	Picture stimuli [N = 226]	24:48	PS
	Zoom	08:02	15
G1F4	Picture stimuli [N = 233]	18:15	PS
	While You Are Sleeping	08:53	16

Table 1.6: Inventory of primary recording sources

#### 1.4.2 Secondary data sources

Secondary data comes from audio recordings gathered in McFarland (2017) from two speakers, and from two internet videos. Table 1.7 shows the breakdown of speaker profiles and collected data for secondary sources.

Speaker ID	Generation	Gender	Languages	Recordings		
				Name	Time (H:MM:SS)	ID
G2F1US [audio only]	2	F	Khmer (N) Mandarin (Pr) English (Pr)	Elicitation [N = 4]	1:33:48	E
				Little Snail	05:12	17
				Smart Fox and Tiger	10:54	18
				Flower	03:22	19
				Little Carp	10:09	20
				The Pear Story	04:44	21
				Picture stimuli (Lau, 2016)	13:24	PS
				Facebook Messenger		E
G2M2US [audio only]	2	M	Khmer (N) Mandarin (F) English (Pr)	Elicitation	20:36	E
				The North Wind and the Sun	00:48	22
				Facebook Messenger		E
G1F5US	1	F	Khmer (N)	Garden tour	11:40	23
G1M2	1 (presumed)	M	Unknown	News interview	05:44	24
G1F6	1 (presumed)	F	Unknown	News interview	05:44	24
G1F7	1 (presumed)	F	Unknown	News interview	05:44	24

Table 1.7: Secondary consultant profiles and recording list

Speakers G2F1US and G2M2US, who are married, are G2 speakers according to the classification in (§1.4.1.2). They were born in Cambodia and grew up speaking Cambodian Teochew in their households and with family over the years. They left Cambodia and immigrated to the United States in the 1980s. These speakers' IDs include US to indicate their current non-residence in Cambodia. They are multilingual, speaking varying degrees of Khmer, English, Mandarin, and Teochew, Now, however, their primary languages are Khmer and English, and their Teochew has experienced some amount of attrition. G2M2US is the nephew of speakers G1M1 and G1F1, and the cousin of G2M1.

Data on Cambodian Teochew in McFarland (2017) was collected via elicitation, story translations, and picture and video descriptions. All interviews were conducted in English. Recordings were done using the Voice Memo App on an Apple iPhone 5S. A 100-word Swadesh (1955) list was used for initial elicitation with G2F1US. Words were read and shown to her in English and Mandarin and she provided the associated word in Teochew. Sentences and phrases were also elicited over three sessions with G2F1US and one session with G2M2US by asking them “How do you say \_\_\_\_ in Teochew?”. Speaker G2F1US provided additional data by looking at picture story apps on an iPad and listening to the Mandarin narration and retelling the story page by page in Teochew. Stories from RyeBooks included *The Flower*, *The Little Carp that Jumped over the Dragon Gate*, *The Little Snail*, *The Smart Fox and the Tiger*.<sup>7</sup> Speaker G2M2US provided an audio recording of Aesop's tale *The North Wind and the Sun* translated from written Mandarin to Teochew. A small amount of more naturalistic data was gathered from speaker G2F1US using picture stimuli from Lau (2016) designed to encourage the production of relative clauses. G2F1US also watched and narrated *The Pear Story* in

---

<sup>7</sup> These story apps are no longer available in the Apple App Store.

Teochew. Finally, outside of the data gathering processes from McFarland (2017) as previously described, both speakers have answered language questions and provided elicited data over Facebook Messenger and in person as needed throughout the duration of the researcher's PhD candidature.

Some additional secondary data used in this dissertation comes from two public videos found on YouTube. One is of speaker G1F5US giving a tour of her garden in California, using Teochew with some Khmer. Speaker G1F5US was raised speaking Cambodian Teochew (and Khmer) in Cambodia, and she has since immigrated to the United States. Thus, her ID is designated with US. Finally, a clip from China's Shantou TV Station 2 (汕头 2) entitled "Chaozhou people in Cambodia" (旅柬埔寨的潮州人) uploaded to YouTube in 2015 includes interviews with three different Cambodian Teochew speakers. Interviews are about speaker background and the proliferation of Teochew opera and culture in Cambodia. Based on the content of the interviews, as well as speaker appearance, the three speakers are presumed to be G1. No information is known for certain about their language background.

In total, secondary sources have provided around three hours of additional data.

### 1.5 Data presentation

While there is no official orthography outside of written Chinese characters, a few different Romanization systems have historically been used for Teochew. The Peh-oē-jī system, or 'Church Romanization', was created in the 19<sup>th</sup> 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*. In 2002, language enthusiasts from the organization Gaginang 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.<sup>8</sup>

Given my speakers' non-usage of Romanization schemes, throughout this dissertation, the collected data on Cambodian Teochew is presented in the International Phonetic Alphabet (IPA). Following Chao (1968), 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. Tone sandhi is represented per convention where 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. 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.

As mentioned above, if the consultants are messaging electronically, they do not use Chinese characters to represent Teochew. However, Chinese characters have been included in this dissertation to provide historical context to the words used in Cambodian Teochew and to increase the research's accessibility among Chinese linguists. Simplified Chinese is used for three reasons. First, the young generation in Cambodia, including the G3 consultants in this study, use Simplified Chinese as that is what is taught in the Chinese schools, whose textbooks and curriculum are heavily influenced by China, with many teachers hailing from the country

---

<sup>8</sup> 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).

(Kyne, 1999; Marks, 2000). Second, characters were primarily sourced from Mogher, a Teochew dialect website out of China that uses Simplified Chinese. Finally, given the current researcher's base in Singapore where Simplified Chinese is used, and that my own language schooling has been exclusively in Simplified Chinese, it was also chosen for personal convenience.

Selecting the appropriate Chinese character proceeded as follows. Mogher was first checked for characters by searching the Teochew pronunciation. The vast majority of the characters used resemble the historical pronunciation found in Mogher (with some regional differences as explored in Chapter 4). There are some cases where the character does not match its historical pronunciation but instead is generally accepted for its meaning. An example of this is 肉 'meat' whose pronunciation should be *nek<sup>5</sup>* but is instead *ba<sup>2</sup>*. In cases where the word was not found in Mogher, dictionaries such as Cai (1991), Fielde (1883), Goh (2017), Goh (2020), TCKnow LLC. (2015),<sup>9</sup> were consulted. The Goh sources include characters for all syllables including loanwords whereas Cai (1991) marks some syllables with no character as □, while TCKnow LLC. (2015) marks them with \*. For example, 'rat' is 獠鼠 in Goh (2020), but □鼠 in Cai (1991) and \*鼠 in TCKnow LLC. (2015). In these cases, the Goh (2020) character was usually chosen and is marked with an asterisk, as in 獠\*鼠. Words borrowed or mixed from English or Khmer are indicated by the use of Khmer or English orthography provided in place of the Chinese character.

Language examples are provided over four lines, as shown in (7): Chinese, Khmer, or English; IPA; interlinear glossing; and English free translation with recording source. Source is

---

<sup>9</sup> Cai (1991), Fielde (1883), and TCKnow LLC. (2015) use Traditional Chinese. Any characters taken from these sources have been converted to Simplified Chinese.

made up of speaker ID, followed by a hyphen and the recording ID. For recording ID, numbers indicate the text number, as found in Tables 1.6 and 1.7, while E represents elicited data and PS is for data from picture stimuli.

(7) 只是个	也是个	Chinese (汉字), English, or Khmer
tɕi <sup>52</sup> kai <sup>55</sup>	mi <sup>75</sup> kai <sup>55</sup>	IPA
PDCO	what	English interlinear gloss
‘What is this?’ [G3F1-E]		‘English translation’ [source]

## 1.6 Organization of the thesis

This dissertation consists of nine chapters including the current chapter, Chapter 1. Chapter 2 covers historical information on the ethnic Chinese in Cambodia and the current sociolinguistic situation for Cambodian Teochews, and makes comparisons to Teochews in other parts of Southeast Asia. Chapter 3 presents a review of the prior literature on Teochew including varieties in China, Cambodia, Indonesia, Malaysia, Singapore, and Thailand.

Moving into the grammar of Cambodian Teochew, Chapter 4 is a traditional analysis of the language’s phonology. While Croft (2001) criticized “reductionist theories of ... phonological representation” (p. 61) and proposed the creation of a Radical Templatic Phonology (Vihman & Croft, 2007), it is not currently a developed theory. For this reason, along with my desire for this grammar to have broader reach and accessibility, I have chosen not to adapt the phonological analysis to the spirit of RCG. However, the remaining parts of the dissertation follow in this framework.

The morphosyntax portion of the grammar is organized into three chapters according to the three propositional act functions of reference, modification, and predication, from Croft

(forthcoming). Thus Chapter 5 covers reference and referent expressions including CT Object Reference Constructions, CT Property Reference Constructions, and CT Action Reference Constructions. Chapter 6 details modification including CT Property, Object, and Action Modification Constructions. Chapter 7 focuses on predication with CT Action, Object, and Property Predication Constructions.

Chapter 8 highlights some CT loanwords from Khmer, as well as their classification and phonological and morphological integration into Cambodian Teochew, while also making comparisons to other Teochew varieties.

Finally, Chapter 9 details the limitations of the current study, ideas for future research, as well as some concluding remarks.

## Chapter 2 – Ethnic Chinese in Cambodia and Beyond<sup>10</sup>

This chapter provides terminology (§2.1) and historical background information (§2.2) on the ethnic Chinese in Cambodia. It also covers the group’s current cultural resurgence (§2.3) and sociolinguistic situation (§2.4) and makes comparisons to Teochews in other parts of Southeast Asia (§2.5). Finally, it showcases some of the Chinese loanwords found in the languages of the region including Khmer and Thai (§2.6). Data for this chapter has been compiled from the literature as well as information gathered from consultant and family interviews, Facebook conversations, and fieldwork observations.

### 2.1 Terminology used for the Cambodian Chinese

There are several terms in the Khmer language to describe ethnic Chinese people.<sup>11</sup> Those who were born in China and emigrated to Cambodia are referred to as *cənc<sup>h</sup>av* (ចិនរើវ, 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 *kooncəvcən* (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

---

<sup>10</sup> Portions of this chapter have been adapted from McFarland (2021).

<sup>11</sup> The terms that follow have been adapted from Edwards (2009) according to their usages by consultants interviewed for this project.

as *cəndaekook* (ចិនដីកែក, ‘mainland Chinese’). Finally, *cəntiəciw* (ចិនទាជីវ, ‘Teochew’) refers specifically to those with Teochew heritage.

As for Chinese/Teochew terms, *hua<sup>11</sup>jiŋ<sup>55</sup>* (华人, ‘Chinese’) can be used for all ethnic Chinese. *hua<sup>11</sup>k<sup>h</sup>iaur<sup>55</sup>* (华侨, ‘overseas Chinese’) historically referred to Chinese citizens abroad. Nowadays it may be used for some speakers who don’t have Chinese nationality. *hua<sup>11</sup>i<sup>24</sup>* (华裔, ‘overseas Chinese’) is a term used for people of Chinese heritage with citizenship from another country. Most Cambodian Chinese are technically *hua<sup>11</sup>i<sup>24</sup>*, though my consultants reported unfamiliarity with the term, or rarity of use. *tuŋ<sup>11</sup>naŋ<sup>55</sup>* (唐人, ‘Chinese’), or people from the Tang (唐) dynasty, is also used by some to refer to all Chinese people, and sometime specifically all overseas Chinese. Some Teochews in Cambodia use *tuŋ<sup>11</sup>naŋ<sup>55</sup>* to refer specifically to Teochew people. This may be due to the group’s current and historic majority in the country. Some of the nuanced differences between the aforementioned 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 *tiə<sup>11</sup>təiur<sup>33</sup>naŋ<sup>55</sup>* (潮州人, ‘Teochew people’) or *ka<sup>33</sup>ki<sup>11</sup>naŋ<sup>55</sup>* (家己人, ‘our own people’). In Cambodia, they may also say they are *kaŋ<sup>24</sup>pəu<sup>33</sup>təe<sup>11</sup>tuŋ<sup>11</sup>naŋ<sup>55</sup>* (柬埔寨唐人, ‘Cambodian Chinese’) or *kaŋ<sup>35</sup>pəu<sup>33</sup>təe<sup>11</sup>hua<sup>11</sup>jiŋ<sup>55</sup>* (柬埔寨华人, ‘Cambodian Chinese’).

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, p. 553; Edwards, 2009, p. 176; Tan, 2006, p. 155-6), while Sino-Cambodian refers to a broader group of ethnic Chinese people with Cambodian citizenship (Verver, 2012, p. 49). Speakers consulted for this project were generally unfamiliar with the Sino- terms. One speaker prefers 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 paper 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 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 (Chan, 2005), 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 (Huiguan) in Phnom Penh in 1814 (S. A. Chen, 2015). Huge waves of Teochew immigration came in the 1930s to the 1960s, mostly from Jieyang, Chaoyang, and Puning, China, 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, p. 199). By 1897, 22,000 of the 50,000 inhabitants (44%) of Phnom Penh were Chinese (Siphath, 2017, p. 185). In Cambodia as a whole, there were reportedly 106,764 Chinese in 1874, representing 11% of the country’s population (Schliesinger, 2011, p. 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, p. 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 dialect 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 (Willmott, 1967, pp. 87-9). 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; Kiernan, 1986, 1990). Finally, when the Vietnamese took over in the 1980s, the Chinese faced continued suppression in retaliation for China's support of the Khmer Rouge (Clayton, 2006; Edwards, 2002; McGee, 2017).

As a result of this decades-long oppression, Cambodia saw a mass emigration of the Chinese population, most of whom would never return.<sup>12</sup> 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; Tea, 2018; Ung, 2000). By 1984, there were reportedly only 61,400 Cambodian Chinese in the country (Siphat, 2017, p. 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, p. 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, 2012b).

---

<sup>12</sup> See Tan (2006), Chapter 6A for information on those who did return.

### 2.3 Current Chinese resurgence

Recent times have seen a resurgence of the Chinese in Cambodia. They have reestablished their strength in business (Filippi, 2013; Marks, 2000; Siphat, 2017; ter Horst, 2008; Verver, 2012), with the help of new migrants from mainland China, Hong Kong, and Taiwan (Chin, 2017; Tan, 2006). 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 dialect associations and social organizations have been resurrected (S. A. Chen, 2015; Marks, 2000; Tan, 2006) and Chinese schools have reopened. The success and prestige of the current Chinese schools has even led to an enrollment of 10-30% of ethnic Khmers due to the perceived increase in business opportunities afforded by speaking Mandarin (Edwards, 2002; Marks, 2000; Tea & Nov 2009; Tan, 2006).

Based on my observations and experiences in Cambodia, evidence of Chinese heritage and influence is everywhere. Visually, a number of signs displaying written Chinese can be found around the city. Signage includes permanent storefronts as well as fixed and moving advertisements. Kasanga's (2012) study on the linguistic landscape of a commercial neighborhood in central Phnom Penh found a number of advertisements displaying Chinese and Khmer. He concluded that "Chinese is experiencing a comeback in view of its historical presence and, to a lesser extent, the growing influence of China in the region" (Kasanga, 2012, p. 565). Other displays of Chinese-ness found around the city include altars with incense and offerings to *ti<sup>11</sup>tɕu<sup>51-25</sup>sij<sup>55</sup>* (地主神, pinyin: dìzhǔ shén), the Landlord God, as well as ancestor shrines, and couplets in Chinese characters posted above and around storefronts and home doorways for good luck. Observed Chinese cultural practices include the tradition of burning of paper money and objects for deceased relatives (see Ann, 2011), and the celebration of holidays

such as Chinese New Year, Hungry Ghost Festival, Mid-Autumn Festival, and Tomb Sweeping Day.

While these displays and celebrations speak to the Chinese influence in Phnom Penh and around the country, they don't necessarily equate to the number of people with Chinese heritage. Instead, many businesspeople have adopted certain Chinese customs despite claiming no Chinese heritage because "[t]he 'Chinese' lifestyle radiated symbols of business success and power" (Dahles & ter Horst, 2012, p. 222). And so, Khmer wholesalers and middlemen in the silk industry may wear traditional Chinese fashion and showcase Chinese paraphernalia (Dahles & ter Horst, 2012; ter Horst, 2008) while some urban, middle-class Khmer might choose Chinese burial customs over traditional Khmer cremation (Davis, 2009, 2013). Even when we look to the people in Cambodia who can claim a Chinese heritage, ethnic belonging and identity are complicated notions, as evidenced by the variety of terminology used to describe the Cambodian Chinese outlined in (§2.1). Edwards (2002) explored this "extraordinary diversity and range of meanings of 'being Chinese' in Cambodia over the past three decades" (p. 255) and in Edwards (2009), she warned of "creat[ing] illusory patterns of ethnicity" (p. 175) by strictly categorizing individuals with Chinese names, who participate in Chinese religious rituals and holidays, as Chinese. Even language isn't a distinguishing factor in a Cambodian Chinese/Teochew identity, as exemplified by an interviewee in Verver (2012) who said "I feel Khmer, but everybody calls me Chinese. I speak the Teochew dialect, but I cannot read or write, and I don't speak Mandarin...My mother and father are from Cambodia, I am from Cambodia" (p. 37). Similar sentiments were expressed by speaker G3F1 in my own research who strongly identifies as Cambodian.

Given the blurred lines between Chinese and Khmer identities explored above, the long

history of intermarriage, historical fear of admitting to a Chinese background, and a general lack of desire to be a statistically unique entity (Tea & Nov, 2009), calculating the number of ethnic Chinese in Cambodia is a challenge, 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) (Siphath, 2017, p. 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, p. 166). X. Chen (2009) and Edwards (2009) reported that 80% of the Chinese in Cambodia are of Teochew origin.

#### 2.4 Sociolinguistic situation of the Teochew in Cambodia

Along with notions of identity, the sociolinguistic situation and language practices of the Cambodian Teochew are also complicated and varied. From the literature, 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 Chinese dialect 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, p. 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 and Koning (2018).

Meanwhile, it was found that language use and choice is situational. While Khmer, Mandarin, and Chinese dialects 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 (S. A. Chen, 2015; Bourgerie, 2017) and often in business (Tea & Nov, 2009; Verver & Koning, 2018), and Mandarin is spoken in school and for association matters (S. A. Chen, 2015, p. 121), in business relations with mainland Chinese and in Southeast Asia (Verver & Koning, 2018), and in communication between different Chinese groups (Filippi, 2010). The current researcher’s observations and interviews similarly reflect these findings. Consultants reported speaking both Teochew and Khmer with family and some friends, Khmer with non-Teochew-speaking individuals, and Mandarin in school or for business. Around the city, I heard conversations in Teochew amongst friends at the gym, and saw business interactions in Teochew at Central, Orussey, and Phsar Toul Tom Pong markets.

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. Several factors are at play. First, not speaking a minority language at home is detrimental to its transmission (Fishman, 1991). Evidence points to the reduced usage of Teochew at home in Cambodia. In addition to the student survey described above, Bourgerie (2017) surveyed a non-student group of 62 small-business and middle class respondents, aged 20-80 years, for comparison. Results of respondents’ home language use are shown in Table 2.1 for the non-student group and in Table 2.2 for the student group.

Language at Home	<b>Khmer Primary</b>	<b>Teochew Primary</b>	<b>Cantonese Primary</b>	<b>Khmer &amp; Teochew</b>	<b>Khmer &amp; Cantonese</b>	<b>No Khmer</b>	<b>Unknown</b>
Percentage	55%	32.5%	12.5%	17.5%	5%	27.5%	5%
Count	22	13	5	7	2	11	2

Table 2.1: Non-student survey: Home language use (Bourgerie, 2017, p. 173)

Language at Home	<b>Khmer Primary</b>	<b>Teochew Primary</b>	<b>Cantonese Primary</b>	<b>Khmer &amp; Teochew</b>	<b>Khmer &amp; Cantonese</b>	<b>No Khmer</b>	<b>Unknown</b>
Percentage	81%	12.9%	2%	34%	4.1%	6.8%	1.4%
Count	119	19	3	50	6	10	2

Table 2.2: Student survey: Home language use (Bourgerie, 2017, p. 173)

The differences between Tables 2.1 and 2.2 showcase the shift away from Teochew in Cambodia. While 32.5% of non-students reported primarily using Teochew at home, only 12.9% of students did so. Conversely, there is an increase of Khmer usage, with 55% of the non-student group reporting primarily speaking that language at home versus 81% of student respondents. More mixing was also reported by the students, with 34% reportedly mixing Khmer and Teochew, versus the non-students at 17.5%. Nyíri (2012) has similarly described a shift from Teochew to Khmer in the home language of young Chinese Cambodians (p. 106). Of the families consulted for this study, it was also noted that children raised in households that primarily spoke Khmer were not likely to speak Teochew proficiently. In contrast, the three G3 consultants reported speaking a mixture of Khmer and Teochew at home, likely the key factor in their families’ continued Teochew transmission.

Institutional factors such as government policy and the language of education have also played a big role in the decline of Teochew in the country. As mentioned in (§2.2), political

policies during the 1970s and 1980s greatly crippled the Cambodian Teochew community's language use. Furthermore, Chinese schools began switching the language of instruction from the dialects to Mandarin when the country established diplomatic relations with China in 1958 (Edwards, 2009, pg. 185; Goh, 2011, p. 11; Willmott, 1967, p. 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, p. 106). This switch to Mandarin in schools and increasing use of the language by Cambodian Chinese was predicated on the belief that speaking it would unite the various Chinese groups in the country (Bourgerie, 2017; Kyne, 1999). With the classroom no longer a domain for Teochew language use, speaking it at home became its primary form of transmission, and it has already been shown that many families have shifted to Khmer at home.

Finally, despite the reported desire to preserve Chinese language and traditions by marrying other ethnic Chinese (Tea & Nov, 2009), exogamy has happened and ultimately led to language shift for the younger Cambodian Teochew generations (Filippi, 2010; Willmott, 2006).

## 2.5 Sociolinguistic situation of the Teochew in Southeast Asia

Similar language shifts are being seen in Teochew communities across Southeast Asia for many of the same reasons. It's been reported that many families have moved away from Teochew in the home domain in Indonesia (Veniranda, 2015), Malaysia (Ong, 2018; Sim, 2012; Wang, 2016), Singapore (Lee, 2015; Li, Saravanan, & Ng, 1997; Ng 1996) and Thailand (Lee, 2014). Interestingly, in Thailand, Lee (2014) found that some study participants were able to learn Teochew from business settings, despite speaking Thai at home (p. 191), though this is very rare.

Institutional factors are also at play in each of these countries. In Singapore, the launch of the government's Speak Mandarin Campaign in 1979 was instrumental in the move away from the dialects (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. Neighboring 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 dialects in Chinese schools in the region, starting in the 1920s (Wang, 2016). 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, p. 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 the Teochew, as well as the addition of Mandarin as an extracurricular activity or obligatory subject in both private and public 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 (Morita, 2007; Chokkajitsumpun, 1998). 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 (Lee, 2015; Li et al., 1997; Ng, 1996; Chokkajitsumpun, 1998; Quek, 2013). For example, in the markets of Johor Bahru, 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, Mandarin usage has increased in an attempt to unify the multiple Chinese groups in the respective countries (Lee, 2015; Li et al., 1997; Ong, 2018; Chokkajitsumpun, 1998; Sim, 2012; Stenberg, 2015). Furthermore, exogamy has changed the speaking practices of Teochews in Indonesia (Stenberg, 2015; Veniranda, 2015), Malaysia (Ong, 2018; Sim, 2012), Singapore (Ng, 1996; Quek, 2013), and Thailand (Lee, 2014).

## 2.6 Chinese/Teochew influence on languages of Southeast Asia

Despite the more recent language shift, it's 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 dissertation is on Cambodian Teochew and Khmer's influence on it, Chinese varieties have also been found to have affected Khmer vocabulary. This is unsurprising given the high percentage of Chinese in Cambodia 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, p. 1), with significant importance and influence, especially in the business sector.

Pou and Jenner (1973) investigated Chinese language 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 Chinese dialects. For example *c<sup>h</sup>á* ‘fry’<sup>13</sup> that became Khmer *chaa* was attributed to Amoy/Hokkien (Pou & Jenner, 1973, p. 47) while *p<sup>h</sup>ũe* ‘blanket’ that became Khmer *phuuəj* was said to come from Teochew (Pou & Jenner, 1973, p. 16). In several cases, the source Chinese was the same in multiple varieties, so the exact donor language could not be determined, such as with *só* ‘lock’. This became Khmer *saa*, and was attributed to Amoy/Hokkien/Cantonese (Pou & Jenner, 1973, p. 83). Pou and Jenner noted that the bulk of items appeared closest to Hokkien (1973, p. 4). This could be because the Hokkiens were reportedly the first Chinese group to settle in Cambodia (Chan, 2005; Siphath, 2017). The vocabulary items were typically new things

---

<sup>13</sup> 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.

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 theater (6%), kinship terms (5%), technological terms (arts and crafts) (4%), administrative and legal (2%), religious (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. 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* dictionary and some were noted to be used only amongst Chinese speakers. According to Edwards (2012a, p. 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 variety of Teochew that have been adopted into Thai, explaining how the word's pronunciation changed in the borrowed Thai form. 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) similarly 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, daily necessities.

Table 2.3 below compares person terminology borrowed from Teochew/Chinese found in Thai, Khmer, and other languages of the region.<sup>14</sup> As mentioned above, exact language origins are difficult to pinpoint. The majority of the words in Table 2.3 were attributed to Teochew but a few were said to come from another Chinese variety.

Teochew/Chinese	English	Khmer	Thai	Other
tsiŋ <sup>33</sup> teŋ <sup>33</sup>   亲丁	a trusted follower	x	tɕi:n <sup>33</sup> teŋ <sup>33</sup>   จีนเต็ง	
tsɔ <sup>32</sup> sũa <sup>33</sup> 座山	a rich man, esp. Chinese	x	tɕaw <sup>41</sup> sua <sup>24</sup>   เจ้าสัว	
t <sup>h</sup> aw <sup>21</sup> kɛ <sup>33</sup>   头家	shopkeeper, tradesman	thaw kaaɛ	thaw <sup>41</sup> kɛ: <sup>22</sup>   ถ้าวแกก	Lao: thâw kɛɛ
t <sup>h</sup> aw <sup>21</sup> kɛ <sup>33</sup> niɔ <sup>55</sup> 头家娘	wife of the boss	x	thaw <sup>41</sup> kɛ: <sup>33</sup> nia <sup>453</sup> ถ้าวเกเนี้ย	
t <sup>h</sup> aw <sup>21</sup> naŋ <sup>55</sup> 头人	a respected, elder Chinese	x	thaw <sup>41</sup> naŋ <sup>453</sup> ถ้าวนัง	
tua <sup>21</sup> puŋ <sup>55</sup>   大肥	to be pot-bellied, fat	x	tua <sup>41</sup> puŋ <sup>453</sup>   ตัวปุย	
tiŋ <sup>21</sup> naŋ <sup>55</sup>   唐人	Chinese person	x	tiŋ <sup>41</sup> naŋ <sup>453</sup>   ตังนัง	
huan <sup>33</sup>   番	foreigner, non-Chinese	x	huan <sup>33</sup>   ฮวน	
sia <sup>11</sup>   舍	a wealthy Chinese	x	sia <sup>22</sup>   เสี่ย	
si <sup>33</sup> tse <sup>51</sup> 师父	form of address for teacher	x	si: <sup>33</sup> tɕe: <sup>453</sup> ซือแป๋	
si <sup>33</sup> hĩa <sup>33</sup> 师兄	form of address for older male classmates	x	si: <sup>33</sup> hĩa <sup>33</sup> ซือเฮี้ย	

<sup>14</sup> Teochew pronunciation and Chinese characters come from Gyarunsut (1983). Note that these pronunciations may be different from those found in the presented data on Cambodian Teochew. Characters have been converted to simplified Chinese. English translations have been adapted from Cooper (2020) and Pou & Jenner (1973). 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. Khmer pronunciation is from Pou & Jenner (1973) unless otherwise noted. Thai script and pronunciations come from Gyarunsut (1983). Words from other languages are from Pou & Jenner (1973) unless otherwise indicated.

sam <sup>33</sup> pɔ <sup>24</sup> koŋ <sup>33</sup> 三保公	a Chinese Buddha image	x	sam <sup>33</sup> pɔ <sup>33</sup> koŋ <sup>33</sup> ซาปออง	
zi <sup>21</sup> pŭa <sup>55</sup>   二盘	a second-level seller	x	ji: <sup>41</sup> pua <sup>453</sup>   ยี่บัว	
sã <sup>33</sup> pŭa <sup>55</sup>   三盘	a third-level seller	x	sa: <sup>33</sup> pua <sup>453</sup>   ซาบัว	
now <sup>33</sup> kŭa <sup>51</sup>   孛仔	child	x	noŋ <sup>33</sup> kia <sup>55</sup>   โนงเกียะ	
kuŋ <sup>33</sup> si <sup>33</sup>   军师	an advisor	x	kun <sup>33</sup> si: <sup>33</sup>   กุนซือ	
hŭɔ <sup>33</sup> koŋ <sup>33</sup>   香公	the keeper of a temple	x	hia <sup>33</sup> koŋ <sup>33</sup>   เฮียกง	
(ʔa <sup>33</sup> )ku <sup>24</sup>   (阿)舅	uncle, mother's brother	ʔaku: <sup>15</sup>	(ʔa: <sup>33</sup> ) ku: <sup>24</sup>   (อา)กู่	
(ʔa <sup>33</sup> )ʔi <sup>55</sup>   (阿)姨	aunt, mother's sister	x	(ʔa: <sup>33</sup> ) ʔi: <sup>453</sup>   (อา)อี	
(ʔa <sup>33</sup> )pe <sup>ʔ22</sup> (阿)伯	uncle, father's brother; old man	(qaa) pèq	(ʔa: <sup>33</sup> ) pe <sup>ʔ55</sup> (อา)แป๊ะ	Lao: qaa pēq
(ʔa <sup>33</sup> )mue <sup>21</sup> (阿)妹	younger sister; Chinese girl	(qaa) múuej	(ʔa: <sup>33</sup> ) muaj <sup>24</sup> (อา)หมวย	Lao: (qaa) mŭej
(ʔa <sup>33</sup> )tse <sup>51</sup>   (阿)姐	older sister	caaɛ	(ʔa: <sup>33</sup> ) tɕe: <sup>453</sup>   (อา)เจ๊	Lao: cee
(ʔa <sup>33</sup> )ti <sup>24</sup>   (阿)弟	younger brother	(qaa) tii	(ʔa: <sup>33</sup> ) ti: <sup>24</sup>   (อา)ตี้	
(ʔa <sup>33</sup> )kow <sup>33</sup> (阿)姑	aunt, father's sister	kòow	(ʔa: <sup>33</sup> ) ko: <sup>33</sup>   (อา)โก	Lao: (qaa) kōo
(ʔa <sup>33</sup> )koŋ <sup>33</sup> (阿)公	grandfather	(qaa) koŋ	(ʔa: <sup>33</sup> ) koŋ <sup>24</sup>   (อา)กง	Malay: əŋkoŋ Tagalog: ingkong
(ʔa <sup>33</sup> ) sim <sup>51</sup> (阿)婶	aunt, wife of father's bro.; Chinese woman	(qaa) sam	(ʔa: <sup>33</sup> ) sim <sup>453</sup> (อา)ซิม	
(ʔa <sup>33</sup> ) so <sup>55</sup>   (阿)嫂	elder sister in law	sao	(ʔa: <sup>33</sup> ) so: <sup>453</sup>   (อา)ซ้อ	Lao: (qaa) sòv
(ʔa <sup>33</sup> ) hŭa <sup>33</sup> (阿)兄	elder brother; Chinese male	(qaa) hŭə	(ʔa: <sup>33</sup> ) hia <sup>33</sup>   (อา)เฮีย	
tia <sup>35</sup>   爹	father	tŭə	tia <sup>22</sup>   เตี่ย	Lao: tiə
tsek <sup>22</sup>   叔	a Chinese man	(qaa) cək	ɕek <sup>55</sup>   เจ๊ก	Lao: cēk
tio <sup>21</sup> tsiw <sup>33</sup>   潮州	Teochew	tŭə ciw ชาฉีว	te: <sup>41</sup> ɕiw <sup>24</sup>   แต่ฉีว	Lao: tēɛ chŭw
haj <sup>24</sup> nam <sup>55</sup>   海南	Hainan, Hainanese	haj nam ไห่ณาส	haj <sup>24</sup> lam <sup>24</sup>   ไหลล่า	
taj <sup>21</sup> koŋ <sup>33</sup>   舵公	helmsman	taj koŋ	taj <sup>41</sup> koŋ <sup>24</sup>   ไต้กง	
tsuŋ <sup>21</sup> tsu <sup>51</sup>   船主	the master of a ship	cən cuu	ɕun <sup>41</sup> ɕu: <sup>453</sup>   จันจู้	
siŋ <sup>33</sup> sɛ <sup>33</sup>   先生	doctor, teacher	sinsaaɛ	sin <sup>33</sup> se: <sup>24</sup>   ซินแส	

Table 2.3: Teochew appellation loanwords in Khmer and Thai

<sup>15</sup> From Huang and Mo (2017, p. 94).

## 2.7 Summary

This chapter has shown that Chinese influence on the languages and cultural practices in Southeast Asian countries like Cambodia, Thailand, and others, is palpable. But unfortunately, it is not enough to stop the shift away from the minority Sinitic languages such as Teochew. Several factors are working against its continued transmission throughout Southeast Asia, including reduced usage of the language at home and school, governmental policies, and exogamy. And it can be hard to reverse these trends. As Quek (2013) found, language shift in Singapore has led to a whole generation who lacks 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 intergenerational transmission.

## Chapter 3 – Prior Teochew Language Research

This chapter looks at the prior research that has been done on the language of the Teochew across the region including in Chaoshan (§3.1) as well as throughout Southeast Asia (§3.2).<sup>16</sup>

### 3.1 Studies on Chaoshan varieties

Much of the prior Teochew research has been done on varieties in the Chaoshan region of China. This section covers the historical work done by missionaries (§3.1.1), as well as more recent linguistic studies on phonetics/phonology (§3.1.2) and grammar (§3.1.3).

#### 3.1.1 Missionary sources

Foreign missionaries to China provided some of the earliest works on Teochew. Fielde (1883) and Goddard (1888) delivered Teochew to English dictionaries arranged alphabetically according to pronunciation, along with the associated Chinese character. Each work also included a table of the 214 radicals used in the characters. Fielde's dictionary is much more detailed than Goddard's, totaling 5442 words, with each entry containing Romanizations of many compound words, phrases, and sentences. Duffus (1883) provided a detailed English to Chinese dictionary. Teochew words are presented in Romanized form and no Chinese characters are used. Entries contained multiple ways of saying each word as well as additional usages, phrases, and sentences. Ashmore's (1884) grammar book was intended to be used by English speakers to learn to speak the Swatow variety and includes sections on nouns,

---

<sup>16</sup> The terminology used in this chapter reflects the terms used by the respective authors who have mostly taken traditional approaches to grammar, and thus the terms are often contrary to the radical construction grammar approach/terms used in the rest of the dissertation. Where possible, future chapters of this dissertation will reference the prior research and attempt to incorporate their findings into a RCG approach.

pronouns, verbs, adverbs, mood, tense, and slang, among other topics. Chinese characters are not used. Dean (1841), Fielde (1878), Giles (1877) and Lim (1886) each provided phrasebooks arranged topically with Romanized pronunciation. All but Giles (1877) included Chinese characters. Contained in each of the aforementioned sources is a pronunciation guide, though they vary greatly in depth. Giles (1877) mentioned only four pronunciation rules and Dean (1841) just covered the vowel sounds. Duffus (1883) and Lim (1886) detailed the pronunciation of vowels, diphthongs, and consonants and the former also elaborated on tones and nasals. Fielde's (1878, 1883) works were the most comprehensive, with each including illustrations of the tones and minimal pair comparisons of tones, aspirates, and nasality. Fielde (1883) also covered tone sandhi. Overall these missionary sources gave insight into the Teochew spoken in the region during the mid-nineteenth century.

### 3.1.2 Recent phonetic/phonological studies

Several Teochew studies have been from phonetic/phonological viewpoints. Karlgren (1915) offered one of the first descriptive studies of the phonological system of Teochew by a Western linguist. Other descriptive studies of Chaoshan sound systems include Lin (1994b) on the Chenghai variety and Zhang (1981) on Chaoyang.

Many researchers looked at the complicated tone sandhi system of Chaoshan (Cheng and Wang, 1977; Zhang, 1979b, 1980; Zhang, 1992). Other works include Wright (1983) who used a metrical approach to tone sandhi in Chinese dialects, including Chaozhou, and Bao (1999), who used generative phonology to analyze tone and tone sandhi patterns in the Jieyang variety to conclude that tonal register is separate from tonal contour.

Several studies were comparative in nature. Lin's (1973) phonetic study of Chaozhou compared it to Mandarin, Cantonese, Wu, Hakka and Amoy, while also making comparisons to

Guang Yun (广韵), a Chinese rime dictionary from the 11<sup>th</sup> century. Lin found many similarities between Teochew and archaic Chinese, as well as differences between Teochew and ancient and contemporary Chinese, ultimately calling Teochew the contemporary Chinese dialect with the longest history. Li (1986) explored phonetic differences such as rhyme, pronunciation, tone, and tone sandhi in eight Chaoshan dialects (Jieyang, Jiexi, Shantou, Chenghai, Raoping, Chaoyang, Huilai, and Puning). Lin (1994a) studied the phonetics of two dialects in Jiexi county, Hepo (Hakka) and Mianhu (Min), and made comparisons to Shantou, Chaozhou, and Mandarin. Lin (1995) examined tone and tone sandhi in Chaoshan varieties. The focus was on Shantou but comparisons were made to six other varieties (Chaozhou, Jieyang, Chaoyang, Raoping, Chenghai, and Nan'ao). The author showed how differences in surface tone and tone sandhi patterns can result in different interpretations. Lin and Chen (1996) compared the sound systems of Shantou, Chaozhou, Jieyang, Chaoyang, Chenghai, and Haifeng. Lin, Lin, and Xu (2005) looked at the Min language of Nan'ao island and performed a comparative study of the phonetic systems of Houzhai town and Yun'ao town. Lin and Chen (2011) compared tone and tone sandhi patterns of the Waisha, Chenghai, and Shantou varieties. Waisha dialect differed from the other two in that its 阳去 yang qu, low tone (pitch 22) changed to falling tone (pitch 42) in sandhi environments, whereas this tone remained stable in the Chenghai and Shantou varieties. Lin (2019) looked at tone and tone sandhi in six Chaoshan varieties (Raopin, Batou, Chengcheng, Chaozhou, Jieyang, and Anbu). The work included detailed graphs and analyses of final and non-final tones in each dialect and showed how final tone may be different than citation tone in some cases. Xu (2013) analyzed Ashmore (1884), Dean (1841), Duffus (1883), Fielde (1878), Fielde (1883), Giles (1877), Lim (1886), and four other dialect records from the 19<sup>th</sup> century to compare and contrast the phonological systems of

the Teochew varieties spoken in Chaoshan during that period. She concluded that there were not actually too many drastic differences amongst the examined sources. Xu (2016) similarly compared historical missionary sources to dialect rhyme books from the early 20<sup>th</sup> century to examine the changes in pronunciation of finals over a 100-year time period, finding three changes: [-n] to [-ŋ]/[-t] to [-k], [a] to [ɛ], and [ŋ] to [uŋ]/[ŋ] to [uŋ]. Wu (2009) looked at the evolution of nasal endings from [-n] to [-ŋ] in several eastern Guangdong Min dialects based on fieldwork in 166 locations around the province. Appendices include brief descriptions of the phonetic systems of 22 locales in eastern Guangdong, including Chao'an, Chaoyang, Chenghai, and Jieyang, among others. Zou (2007) also looked at word endings, focusing on the geographical distribution of [-m] and [-p] across Chenghai. The author found that in some areas, these endings have been changed to [-ŋ] and [-ʔ/-k], respectively, and concluded that he expects to see further gradual reduction of [-m] and [-p] endings in the region. Xu (2009) found that the reduction in [-m] and [-p] endings was not unique to Chenghai, similarly finding the sound change in the dialects of Chaozhou city.

Finally, some studies were more phonetic in nature. Cun (2009) used instrumental methodology to examine the phonetic characteristic of implosives in five Sinitic languages, including Chaozhou Min, while Cun (2010) provided a phonetic explanation for the sound change from voiceless stops to implosives in Chinese dialects, based on data from four Chaozhou speakers. Hong (2013) provided an in depth look at the phonetics of the Teochew of 12 speakers from Chao'an, performing acoustic analyses of implosives, oral and nasal vowels, the glottal stop, phonation, and tone. He also solicited data from 26 speakers from other areas for a comparison of tone in the different varieties and found that speakers from western Chaoshan tended to have more falling tones than those from the east. Li (2014) acoustically

analyzed the tone of four speakers of the Shantou dialect, measuring fundamental frequency in Praat.

### 3.1.3 Recent grammatical studies

Some studies explored specific grammatical constructions in Chaoshan varieties. On Chaoyang, Zhang (1979a) looked at ten different reduplication patterns, while Zhang (1982) focused on two reduplication patterns used in onomatopoeias. Li and Weng (1987) studied comparative and superlative constructions in Chaoshan. J. Chen (1987) looked at the semantic properties, usages, and features of particles 哩 *li* and 罗 *lo*. Lin (1991) briefly covered several morphological processes in Chaoshan such as affixation, post-nominal modification, and reduplication in verbs, nouns, adjectives, and onomatopoeias. He also highlighted some characteristics of personal pronouns and demonstratives, and negation strategies. Lin (1992) looked at properties and usages of function words including prepositions, conjunctions, auxiliaries, modal particles, and interjections. Shi (1999) covered standard and pragmatic usages and etymology of personal pronouns, as well as meanings and usages of demonstrative and interrogative pronouns in Shantou. Shi's (1996) more detailed work additionally explored polar interrogatives, sentence patterns, aspect, reduplication in verb phrases, grammaticalization of 𠵼 *tã*<sup>213</sup> 'to say' into a complementizer, and usages and etymology of 了 *liau*<sup>53</sup> 'to finish'. Zhuang (2001) studied several negative constructions including those with 唔 *m*<sup>24</sup>, 无 *bo*<sup>55</sup>, 𠵼 *boi*<sup>24</sup> 'will not', 勿<sup>17</sup> *mar*<sup>21</sup> 'not want', 免 *meŋ*<sup>52</sup> 'no need', and 𠵼 *mo*<sup>52</sup> 'not good', while Zhang (1996) focused on the negative word 𠵼 *boi*<sup>24</sup>, and Zhang (1999) explored negation strategies with 未 *bue*<sup>11</sup>. Xu (2005) looked at nine negative markers in the Jieyang dialect. Matthews, Xu,

---

<sup>17</sup> 勿 expresses 唔爱 (Zhuang, 2001, p. 49).

and Yip (2005) studied passive and unaccusative constructions in Jieyang. While it is common for Sinitic languages to use ‘give’ morphology in passive sentences (Yap & Iwasaki, 2003), Teochew presented some divergences from other Sinitic languages. Matthews, Xu, and Yip found the agent to be obligatory in passive constructions in Jieyang. Furthermore, seemingly unique to Southern Min dialects was that the same morphology (‘give’ + ‘3SG’) was used in intransitive, specifically unaccusative, predicates such as ‘The flower has died’. The researchers argued for a case of overt marking of unaccusativity. Matthews and Yip (2008) looked further at these constructions, this time focusing on the Chaoyang variety, while also exploring pretransitive<sup>18</sup> constructions. Chaoyang Teochew has a variety of pretransitive markers leading to frequent surfacing of the object before the verb. These pretransitive constructions are also explored in C. Chen (1996). Xu and Matthews (2013) again looked at ‘give’ morphology in Chaoyang and Jieyang varieties, determining that its grammaticalization pathway went from permissive, to sociative causative, to passive, to unaccusative usages. Finally, Xu and Matthews (2011) explored the polyfunctional morpheme 𠵿 *kai* in the Jieyang variety and found it to be used as a nominal classifier, as well as in genitive, possessive, pronominal, relative clause, nominalizer, and sentence-final stance-marking functions.

Cai’s (1991) and Li and Lin’s (1992) etymological works are well known in the field. Cai’s dictionary first presented the vocabulary of Chaoshan arranged alphabetically by pronunciation and then highlighted vocabulary items that differ from Mandarin arranged topically. Li and Lin showcased Chaoshan lexicology on topics such as human physiology, kitchenware, daily wear, animals, plants, actions, and more.

---

<sup>18</sup> “Pretransitive” is a term coined by Chao (1968, p. 342) referring to constructions where a marked object appears before the verb, often referred to as ‘the *ba* construction’ in Mandarin.

For more comprehensive sources, Li (1959) and Xu (2007) provided full and partial, respectively, grammars of the Jieyang variety and highlighted some of the ways in which Teochew differed from Mandarin and other Sinitic languages. Lin (1996) looked at the Chenghai variety, also making comparisons to Mandarin and Middle Chinese. Lin's (2015) diachronic study of Chaoshan includes topics such as morphology, negation, pronouns, demonstratives, adverbs and quantifiers, as well as comparisons to Hokkien and the historical influence of Cantonese.

### 3.2 Studies on Southeast Asian varieties

Some linguistic studies have been done on Teochew varieties in Southeast Asia including in Cambodia (§3.2.1), Indonesia (§3.2.2), Malaysia (§3.2.3), Singapore (§3.2.4), and Thailand (§3.2.5).

#### 3.2.1 Cambodia

The work done on the Teochew of the country of focus in this dissertation is extremely limited. X. Chen (2009) explored the vernacular reading of Chinese texts by the Teochew in Cambodia, Laos, and Vietnam and included a brief outline of the phonology of each variety. X Chen (2014) looked deeper at seven Southeast Asian Teochew varieties of Phnom Penh Cambodia, Pontianak Indonesia, Vientiane Laos, Johor Bahru Malaysia, Singapore, Bangkok Thailand, and Ho Chi Minh Vietnam. She covered each language's phonology, loanwords, and vocabulary differences, as well as highlighted some key grammatical behaviors and divergences amongst the varieties. It is a great comparative study, though she only worked with a few speakers from each variety so it's hard to tell how widespread her conclusions are. For example, her speakers of Cambodian Teochew differed from the speakers in this study in

certain areas of pronunciation, vocabulary, and grammar which will be highlighted at later points in this paper.

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.

### 3.2.2 Indonesia

On varieties spoken in Indonesia, Peng's (2012) doctoral dissertation studied syntactic constructions in Jambi Teochew and Pontianak Teochew that 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). As previously mentioned, X Chen (2014) included Pontianak Teochew in her comparative study of Teochew varieties of Southeast Asia.

### 3.2.3 Malaysia

X. Chen (2003) studied Malaysian varieties of three Sinitic languages, one of which was Johor Bahru Teochew. The work is valuable and extensive, covering the variety's sound

system, grammar, and vocabulary. The latter section compared words used in Johor Bahru versus in Guangdong, showing the influence of Mandarin, Malay, and English in the Malaysian variety. Chen highlighted 98 Malay and 130 English loanwords in Johor Bahru Teochew. Also included is a chart of 3,800 Chinese characters and their readings in each of the three languages of focus, and a glossary with over 2,400 entries including 60 sentences in full IPA.

Khoo (2017a) analyzed the phonology of the Teochew variety spoken in the Tanjung Sepat village of Selangor, Malaysia, and made comparisons to Guangdong Waisha, from where the inhabitants originated. Data from surveys of two speakers from the Malaysian village was compared to Lin (1996). Khoo found that the overall consonant inventory is the same in the two varieties, though there was some character by character variation in initial consonant pronunciation such as in 乐 ‘music’ *gau<sup>ʔ</sup>* versus Lin’s *ŋau<sup>ʔ</sup>*. Other differences are highlighted in the summary of Khoo (2017c) below.

Khoo (2017b) studied tonal differences of three adjacent fishing villages, Kuala Kurau (KK), Tanjung Piandang (TP), and Sungai Udang (SU), on the border of Perak State and Penang State in Malaysia. Khoo also compared these varieties to that spoken in Puning, China, where the majority of the families in these villages originated. The author found that the 阴平 yin ping, mid level tone, was pronounced with pitch 35 in Puning, TP and SU, and with pitch 33 in KK. Furthermore, Khoo showed that the 阴上 yin shang, high falling tone, varied in sandhi environments. All three Malaysian varieties pronounced that tone with pitch 21 in sandhi, while in Puning it was pronounced with pitch 13. Thus, despite their same origins, the differences in tonal pronunciation between the three villages and their dialect of origin showed the variation that can happen in overseas Teochew communities.

Khoo (2017c) similarly explored phonetic variation in five fishing villages in the

Selangor coastal region of Malaysia, whose residents originated from the Waisha village in Chenghai over 100 years ago. The author used data from Khoo (2017a) and additionally surveyed four male speakers aged 72-90, one from each of the other villages, and made comparisons to Guangdong Chenghai, based on data from Lin (1996) and Lin and Chen (2011). While Guangdong Chenghai has final [-iou], contrasted to Shantou's [-iau] (Lin, 1996, p. 4), the villages in Selangor were found to have [-iau], such as in 条 'strip', pronounced *tiau*<sup>55</sup> in all five Malaysian varieties (Khoo, 2017c). There was also some variation between [-iau]/[-io] as in 招 'summon', pronounced *tsiau*<sup>33</sup> in three Selangor varieties and *tsio*<sup>33</sup> in two. Furthermore, while Lin and Chen (2011) found both the 阳上 yang shang, mid rising,<sup>19</sup> and 阳去 yang qu, low level, tones to be pitch 42 in sandhi environments in Guangdong Chenghai, Khoo found these tone pitches to be 21 and 11, respectively, for all five villages. When comparing the Malaysian varieties, Khoo found that some words alternated in pronunciation between [-ue]/[-oi]. For example, 八 'eight' was pronounced *poi*<sup>l</sup> in three varieties and *pue*<sup>l</sup> in the other two. There were also word onset alternations between [n]/[l] and [m]/[ŋ] amongst the different varieties. For example, 暖 'warm' was pronounced *nuaj*<sup>52</sup> in two varieties and *luaj*<sup>52</sup> in the other three while 危 'danger' was pronounced *mu*<sup>55</sup> in one variety and *ŋũ*<sup>55</sup> in the other four. Additionally, the article looked at vowel nasalization in 30 characters and showed that while certain items, such as 耳 'ear'<sup>20</sup> and 虎 'tiger', maintained their historic vowel nasalization across the five dialects, other characters experienced varying degrees of reduced nasalization across the varieties. For example, nasalization is maintained in four dialects for 第 'ordinal

---

<sup>19</sup> See Chapter 4 for more details on naming conventions used for tones in Teochew.

<sup>20</sup> Pronunciation was not provided for the 30 characters beyond whether or not they contained a nasalized vowel.

prefix’, in three for 支 ‘support’, and in one for 毀 ‘destroy’.<sup>21</sup> The total number of characters with vowel nasalization varied between 16 to 21 out of 30 for all the varieties. Khoo stated that the reduction in vowel nasalization was especially prevalent among young Teochew people in Malaysia (2017c, p. 21). Interestingly, there were words that were nasalized in the Selangor villages’ languages, though they were not in their original pronunciation, such as 移 ‘move’ which was nasalized in four varieties and 架 ‘frame’, nasalized in one.

Finally, Khoo (2018) covered dialect mixing and a unique tone phenomenon of a Teochew fishing village in Parit Jawa in Johor Malaysia, whose descendants hailed from Chao’an, Chenghai, and Jieyang over 100 years ago. Based on surveys from six speakers aged 20-70, originating from the different sub-dialect regions (one from Chao’an, three from Chenghai, and two from Jieyang), the author found that polysyllable words ending in the 阴上 yin shang tone varied in their tonal pronunciation depending on the prior syllables. When a 阳上 yin shang syllable followed one in any of the 阳上 yin shang, 阴去 yin qu, 阴入 yin ru tones in their sandhi form (pitches 35, 53, 5, respectively), it was pronounced as pitch 21. When it followed a word in any other tone, it was pronounced with pitch 53. Monosyllable words in the 阴上 yin shang tone were always pronounced 53. This phenomenon was found in all six speakers, regardless of age or sub-dialect origin (Khoo, 2017c). According to Lin (1995), words in the 阳上 yin shang tone in Chao’an and Chenghai were pronounced with pitch 53 while in Jieyang they were pitch 31. Thus, Khoo hypothesized that the 阴上 yin shang tonal variation has arisen in the Parit Jawa village due to the historic mixing of the three dialects of origin.

---

<sup>21</sup> Mogher also notes nasalization for these three terms, though 毀 ‘destroy’ is apparently pronounced as *hur*<sup>53</sup> in Shantou (versus *hur*<sup>53</sup> in four other varieties).

### 3.2.4 Singapore

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. X Chen's (2014) comparative work also included a great basic overview of the Singapore Teochew variety. Native speaker Goh Eng Choon has made considerable efforts to document and describe Teochew as it is used in Singapore. 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. Goh's (2020) updated dictionary contained over 7,000 entries, each with full Chaozhou Pinyin and tone sandhi, which were regrettably lacking in the first edition. Goh (2018) also produced a learner's guide to Singapore Teochew with 60 lessons and eleven folk rhymes, with full pronunciation of all words including tone sandhi. Li (1991) outlined the sound system of Singapore Teochew and presented a list of 100+ loanwords from English and Malay. Additionally, Li gave over 50 "special words" that arose from influence from Hokkien and other Sinitic languages spoken in Singapore. Zeng (2003) compared the vocabulary of Singapore Teochew to that of Shantou through interviews with people from both locales. He found that 681/737 (92.4%) vocabulary items were synonymous in the two dialects, and attributed most of the differences to borrowings from Malay and English. Based on an 80-minute corpus from eight speakers, Lee (2003) reported on language mixing of English, Mandarin, and Cantonese and language borrowing from Hokkien and Malay in Singapore Teochew. The 125 English words found in

the corpus were primarily nouns including locales like *DBS bank*, *church*, and *canteen*, and technology like *email*, *air-con*, and *handphone*. Verbs used included *start*, *exercise*, and *travel*, while the adjectives *happy*, *tired*, and *lucky*, and the conjunctions *then*, *before*, and *because* were also found. The 18 Mandarin words were vocabulary used in formal schooling like 谢师宴 *xièshīyàn* ‘teacher appreciation banquet’ or those without a Teochew equivalent like 搞错 *gǎo cuò* ‘to make a mistake’. The English and Mandarin words were mixed in to the speakers’ Teochew, maintaining their original form. This contrasts with language borrowing, in which foreign words have been integrated into Singapore Teochew. Lee found eight instances of borrowing from Hokkien, such as 淡薄 ‘a little’, used instead of Teochew’s own 滴仔 ‘a little’. There were 20 borrowings from Malay including 巴刹 *pa<sup>33</sup>sak<sup>5</sup>* ‘market’ from *pasar* and 舒甲 *su<sup>33</sup>ka<sup>2</sup>* ‘to like’ from *suka*.<sup>22</sup> Lee’s work offered good insight into communication strategies of an overseas Teochew community. Finally, 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.

### 3.2.5 Thailand

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

---

<sup>22</sup> Malay and Teochew pronunciations are from Goh (2020), as Lee (2003) did not include any Romanization.

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.

Similar to her work in Malaysia, X. Chen (2019) studied Chinese languages spoken in Thailand, including three Teochew varieties spoken in Bangkok, Chiang Mai, and Hat Yai (as well as Cantonese and Hakka varieties). This study includes an overview of the sound system and grammatical features of each variety as well as vocabulary comparison in a glossary of over 2400 items including 60 full sentences.

### 3.3 Summary

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 on the overseas diaspora such as localized loanwords (X Chen, 2014; Goh, 2017, 2020; Lee, 2003; Li, 1991; Lin, 2006; Low, 2014; Zeng, 2003) and varietal phonological (X. Chen, 2009, 2014, 2019; Khoo, 2017a, 2017b; 2018; Pan, 2000) and grammatical (Peng, 2011, 2012; Pan, 2000) divergences. This dissertation aims to bridge some of this gap by contributing novel work to the fields of Sinitic and Cambodian linguistics, as well as to contact linguistics, hopefully shedding light on the behavior of languages in contact environments.

## Chapter 4 – Phonology

As mentioned in §2.2, Cambodian Teochew families originated from different parts of Chaoshan, China, and thus brought with them differing pronunciations, especially in regards to the vowels and tones of the language. And since CT has no official status in Cambodia and is not taught in schools, these differing dialects have not merged into any one standard ‘Cambodian Teochew’ pronunciation. Thus, the sound system presented in this chapter is based on the most common pronunciations of the speakers interviewed for this project. Any differences shared by multiple speakers that point to more than individual variation are marked as such.

This chapter covers consonants in §4.1 and vowels in §4.2 including monophthongs (§4.2.1), diphthongs (§4.2.2), and triphthongs (§4.2.3). The tone system is explored in §4.3, with an overview of CT’s citation (§4.3.1) and tone sandhi patterns (§4.3.2). Syllable structure is explained in §4.4. Finally, a summary with comparisons to other Teochew varieties is presented in §4.5.

### 4.1 Consonants

Cambodian Teochew has 18 consonant phonemes which are arranged according to their manner and place of articulation in Table 4.1.

	Bilabial	Alveolar	Alveolo-Palatal	Velar	Glottal
<b>Stop</b>					
voiceless unaspirated	p	t		k	ʔ
voiceless aspirated	p <sup>h</sup>	t <sup>h</sup>		k <sup>h</sup>	
voiced	b			g	
<b>Affricate</b>					
voiceless unaspirated			tʃ		
voiceless aspirated			tʃ <sup>h</sup>		
voiced			dʒ		
<b>Fricative</b>		s			h
<b>Nasal</b>	m	n		ŋ	
<b>Approximant</b>					
lateral		l			

Table 4.1: Consonant phonemes in Cambodian Teochew

All consonants are phonemic word initially with the exception of the glottal stop, ʔ. This sound occurs in free variation before word-initial vowels, a common phenomenon in world languages (Garellek, 2013).

伊 [ʔi<sup>33</sup>] ~ [i<sup>33</sup>] ‘3SG’

鞋 [ʔɔi<sup>55</sup>] ~ [ɔi<sup>55</sup>] ‘shoe’

The following minimal and near minimal pairs and triplets showcase the contrasts in stops, fricatives, and affricates in the language.

/p/, /p<sup>h</sup>/ and /b/:

飞 /puε<sup>55</sup>/ ‘to fly’

皮 /p<sup>h</sup>ue<sup>55</sup>/ ‘skin’

尾 /buε<sup>52</sup>/ ‘tail’

北 /pak<sup>2</sup>/ ‘north’

打 /p<sup>h</sup>aŋ<sup>2</sup>/ ‘to hit’

肉 /baŋ<sup>2</sup>/ ‘meat’

/t/ and /t<sup>h</sup>/:

长 /tuŋ<sup>55</sup>/ ‘long’

糖 /t<sup>h</sup>uŋ<sup>55</sup>/ ‘sugar’

图 /tɔu<sup>55</sup>/ ‘image’

涂 /t<sup>h</sup>ɔu<sup>55</sup>/ ‘dirt’

/k/, /k<sup>h</sup>/ and /g/:

旧 /ku<sup>11</sup>/ ‘old’

跼 /k<sup>h</sup>u<sup>55</sup>/ ‘to slap’

牛 /gu<sup>55</sup>/ ‘cow’

鸡 /kɔi<sup>33</sup>/ ‘chicken’

溪 /k<sup>h</sup>ɔi<sup>33</sup>/ ‘river’

鹅 /gɔ<sup>55</sup>/ ‘goose’

/s/ and /h/

孙 /suŋ<sup>33</sup>/ ‘grandchildren’

薰 /huŋ<sup>33</sup>/ ‘tobacco’

嫂 /sɔ<sup>52</sup>/ ‘older brother’s wife’

好 /hɔ<sup>52</sup>/ ‘good’

/tɕ/ and /tɕ<sup>h</sup>/:

酒 /tɕiu<sup>52</sup>/ ‘alcohol’

手 /tɕ<sup>h</sup>iu<sup>52</sup>/ ‘hand’

走 /tɕau<sup>52</sup>/ ‘to run’

草 /tɕ<sup>h</sup>au<sup>52</sup>/ ‘grass’

/tɕ/ and /dʒ/:

蛇 /tɕua<sup>55</sup>/ ‘snake’

热 /dʒua<sup>75</sup>/ ‘hot’

鸟 /tɕiau<sup>52</sup>/ ‘bird’

趯 /dʒiau<sup>11</sup>/ ‘to chase off’

The alveolo-palatal affricate phonemes (/tɕ/, /tɕ<sup>h</sup>/, and /dʒ/) may be a differentiating aspect of the CT phonology as compared to other Teochew varieties. According to the literature, the voiceless affricates are pronounced as /ts/ and /ts<sup>h</sup>/ respectively in Jambi and Pontianak Teochew (Peng, 2012), Jieyang Chaozhou (Xu, 2007), Johor Bahru Teochew (X. Chen, 2003) and other Malaysian varieties (Khoo, 2017c), Singapore Teochew (Low, 2014; Yeo, 2011), and Thai varieties (X. Chen, 2019; Eiam-pailin, 2004; Phadungsrisavas, 2008). While Hong (2013, p. 15) and Xu (2013, p. 228) have noted the existence of these sounds allophonically before front vowels /e/ and /i/ in Chaoshan varieties, CT speakers in this study use /tɕ/ and /tɕ<sup>h</sup>/ in front of all vowels in words such as 多 /tɕɔi<sup>11</sup>/ ‘many’, 水 /tɕui<sup>52</sup>/ ‘water’, 书

/tɕu<sup>33</sup>/ ‘book’, and /tɕ<sup>h</sup>a<sup>52</sup>/ ‘to fry’. Figure 4.1 shows spectrograms from speaker G1F4 (left) and a Singapore Teochew (ST) speaker (Learn Dialect in Singapore, 2020) (right) of the word 多 ‘many’, pronounced tɕoi<sup>11</sup> in Cambodian Teochew and tsoi<sup>11</sup> in Singapore Teochew.

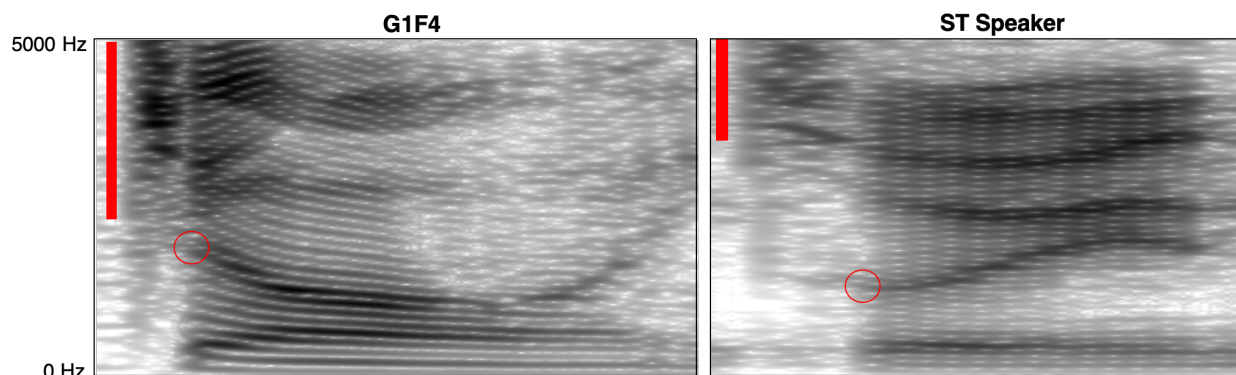


Figure 4.1: Spectrograms comparing CT speaker’s (G1F4) 多 [tɕoi<sup>11</sup>] ‘many’ to a ST speaker’s 多 [tsoi<sup>11</sup>] ‘many’

In Figure 4.1, the differing frequency of the energy in the affricates (te vs. ts) is indicated by the red bar, showing more energy at a lower frequency in the /tɕ/. Additionally, the formant values as the sound transitions into the vowel, as indicated by the red circle, are different. In the CT word, the first formant is low and the second formant is high due to the front and high position of the tongue in the mouth (Ladefoged & Disner, 2012, p. 169), while the second formant is lower in the ST word due to a lower tongue position in the /ts/.

As for the voiced affricate, /dz/, there is variation as to the pronunciation of this sound in the varieties of the diaspora, according to the literature. X. Chen (2003, 2009, 2014, 2019) calls it a /z/ in the Cambodian, Laotian, Malaysian, Singapore, Thai, and Vietnamese varieties, while Yeo (2011) and Low (2014) call it a voice alveolar affricate /ɖ/ in Singapore Teochew and X. Chen (2014) and Xu (2007) similarly have a voiced alveolar affricate /ɖ/ for Pontianak and Jieyang varieties respectively. According to Hong (2013), words with this sound are

pronounced with an alveolar approximant [ʎ] in Chaozhou. Similar to the voiceless affricates, Xu (2013, p. 229) notes that phoneme /dʒ/ was pronounced as [dʒ] before front vowels in the Chaozhou of the nineteenth century. In Cambodian Teochew, at least for the primary G1 speakers, this sound is pronounced as an affricate, with the tongue position close to the palate in all words including in 十二 /tɕap<sup>5-2</sup>dʒi<sup>24</sup>/ ‘twelve’, 日 /dʒik<sup>5</sup>/ ‘sun’, and 入 /dʒip<sup>5</sup>/ ‘to enter’, but also in 热 /dʒua<sup>ʔ5</sup>/ ‘hot’. A spectrogram of G1F4’s 热 /dʒua<sup>ʔ5</sup>/ ‘hot’ is shown in Figure 4.2.

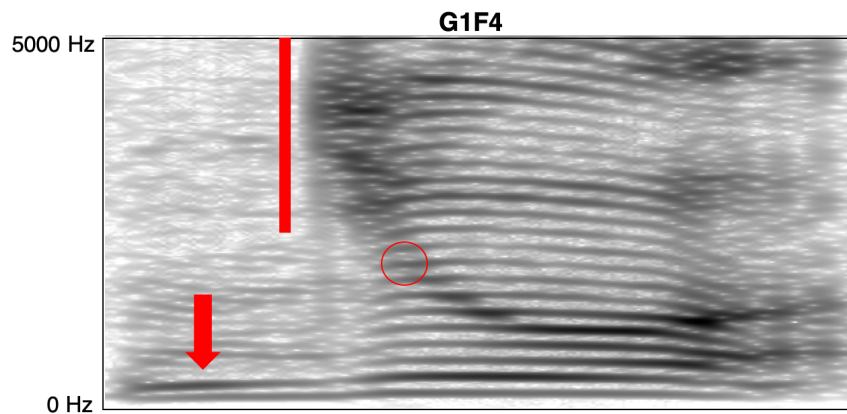


Figure 4.2: Spectrogram of G1F4 热 /dʒua<sup>ʔ5</sup>/ ‘hot’

In the initial sound in Figure 4.2, the prevoicing can be seen in the dark bar at the bottom indicated with the red arrow, while the energy of the fricative is indicated by the red bar, and the fall in the second formant from the high tongue position of the *ʒ* is indicated by the red circle. While G1F4 and the other G1 speakers say /dʒ/, the younger speakers (G2 and G3) in this study seem to have lost the stop part of the affricate and instead pronounce the sound as [ʒ]. This deaffrication may have occurred as a simplification of the phonology, potentially due to the low frequency of words with this sound. This is also consistent with the phonological process of lenition (Hock, 1991).

The near minimal triplets below demonstrate the contrast in nasal initial consonants in Cambodian Teochew.

/n/ vs /ŋ/ vs /m/:

二 <sup>23</sup> /nɔ <sup>24</sup> / ‘two’	俺 /nan <sup>52</sup> / ‘we’
遇 /ŋɔ <sup>24</sup> / ‘to meet’	凝 /ŋan <sup>55</sup> / ‘cold’
毛 /mɔ <sup>55</sup> / ‘hair’	蠓 /maj <sup>52</sup> / ‘mosquito’

Finally, the following near minimal set illustrates the contrast between the five alveolar consonants in word-initial position.

踏 /ta<sup>5</sup>/ ‘to ride (a bike)’  
 读 /t<sup>h</sup>ak<sup>5</sup>/ ‘to study’  
 虱 /sak<sup>2</sup>/ ‘flea’  
 搨 /nak<sup>5</sup>/ ‘to grab’  
 六 /lak<sup>5</sup>/ ‘six’

In word-final position, the following minimal and near minimal pairs illustrate the contrast between /m/ and /ŋ/. The third nasal, /n/, does not occur in this position. Words with historic /n/ codas instead are pronounced with contrastive vowel nasalization (see §4.2.1).

/m/ and /ŋ/

暗 /am <sup>21</sup> / ‘night’	心 /sim <sup>33</sup> / ‘heart’
翁 /aŋ <sup>33</sup> / ‘husband’	新 /siŋ <sup>33</sup> / ‘new’ <sup>24</sup>

<sup>23</sup> This character has two pronunciations depending on its usage, according to Mogher. When used as a single numeral, 二 is pronounced /nɔ<sup>24</sup>/ ‘two’. In larger numbers, it is pronounced /dʒi<sup>24</sup>/ in the tens and ones digits as in 十二 /tɕap<sup>5-2</sup>dʒi<sup>24</sup>/ ‘twelve’ and 二十 /dʒi<sup>24-11</sup>tɕap<sup>5</sup>/ ‘twenty’. 二千 /nɔ<sup>24-11</sup>tɕ<sup>h</sup>ai<sup>33</sup>/ is ‘two thousand’.

<sup>24</sup> This is the pronunciation by G3F2, G3F3, and G1F4. Speakers G1M1, G1F3 and G3F1 say 新 /seŋ<sup>33</sup>/ ‘new’. According to Mogher, seŋ<sup>33</sup> is the 揭阳 Jieyang pronunciation.

The below near minimal triplets illustrate the phonemic difference between glottal, velar, and bilabial voiceless unaspirated stops in syllable-final position.

/p/, /k/ and /ʔ/:

蛤婆 /kap<sup>5</sup>pɔ<sup>55</sup>/ ‘toad’

盒 /ap<sup>5</sup>/ ‘box’

角 /kak<sup>2</sup>/ ‘horn’

目 /mak<sup>5</sup>/ ‘eye’

合 /kaʔ<sup>2</sup>/ ‘with/and’

鸭 /aʔ<sup>2</sup>/ ‘duck’

## 4.2 Vowels

### 4.2.1 Monophthongs

Cambodian Teochew has six distinctive oral vowels /i, e, u, a, ɔ/ and three distinctive nasal vowels /ĩ, ẽ, ã/. These vowels are arranged in Table 4.2.

	Front	Central	Back
Close	i ĩ	u	ɯ
Mid	e ẽ		ɔ
Open		a ã	

Table 4.2: Vowel phonemes in Cambodian Teochew

The following minimal pairs showcase the phonemic contrasts between phonetically similar vowels.

/i/ and /e/:

只 /tɕi<sup>52</sup>/ ‘this’

姐 /tɕe<sup>52</sup>/ ‘older sister’

/e/ and /u/:

虾 /he<sup>55</sup>/ ‘shrimp’

鱼 /hu<sup>55</sup>/ ‘fish’

/u/ and /a/:

书 /tɕu<sup>33</sup>/ ‘book’

昨 /tɕa<sup>33</sup>/ ‘yesterday’

/u/ and /ɔ/:

牛 /gu<sup>55</sup>/ ‘cow’

鹅 /gɔ<sup>55</sup>/ ‘goose’

Formant frequencies are used to characterize vowel quality. The first formant ( $F_1$ ) relates to vowel height (or close/mid/open) while the second formant ( $F_2$ ) relates to the frontness or backness of the vowel. Plotting vowels in an  $F_1 \times F_2$  graph shows the acoustic space of vowels in a language. Figure 4.3 showcases the six oral vowels plotted in a vowel space ( $F_1 \times F_2$ ) for eight primary speakers<sup>25</sup>. Vowels were analyzed from the picture stimuli data from all speakers except G1M1, where the Swadesh elicitation data was used. Praat software was used to measure vowels acoustically. Oral monophthongs were manually labelled in an interval tier in a TextGrid. Each vowel has at least five tokens. Then the `extract_formant_data` script from Ring (2017) was used to pull duration, pitch, formant frequency, and intensity values from the midpoints of the labeled vowels. Finally, vowels were plotted using Ring’s (2017) `draw_formant_plot` script with the already extracted formant data. The resulting graphs are presented in Figure 4.3. Ellipses represent standard deviation of one for each vowel.

---

<sup>25</sup> Due to feedback from the lapel microphone, picture stimuli data from speaker G3F2 is not suitable for phonetic analysis.

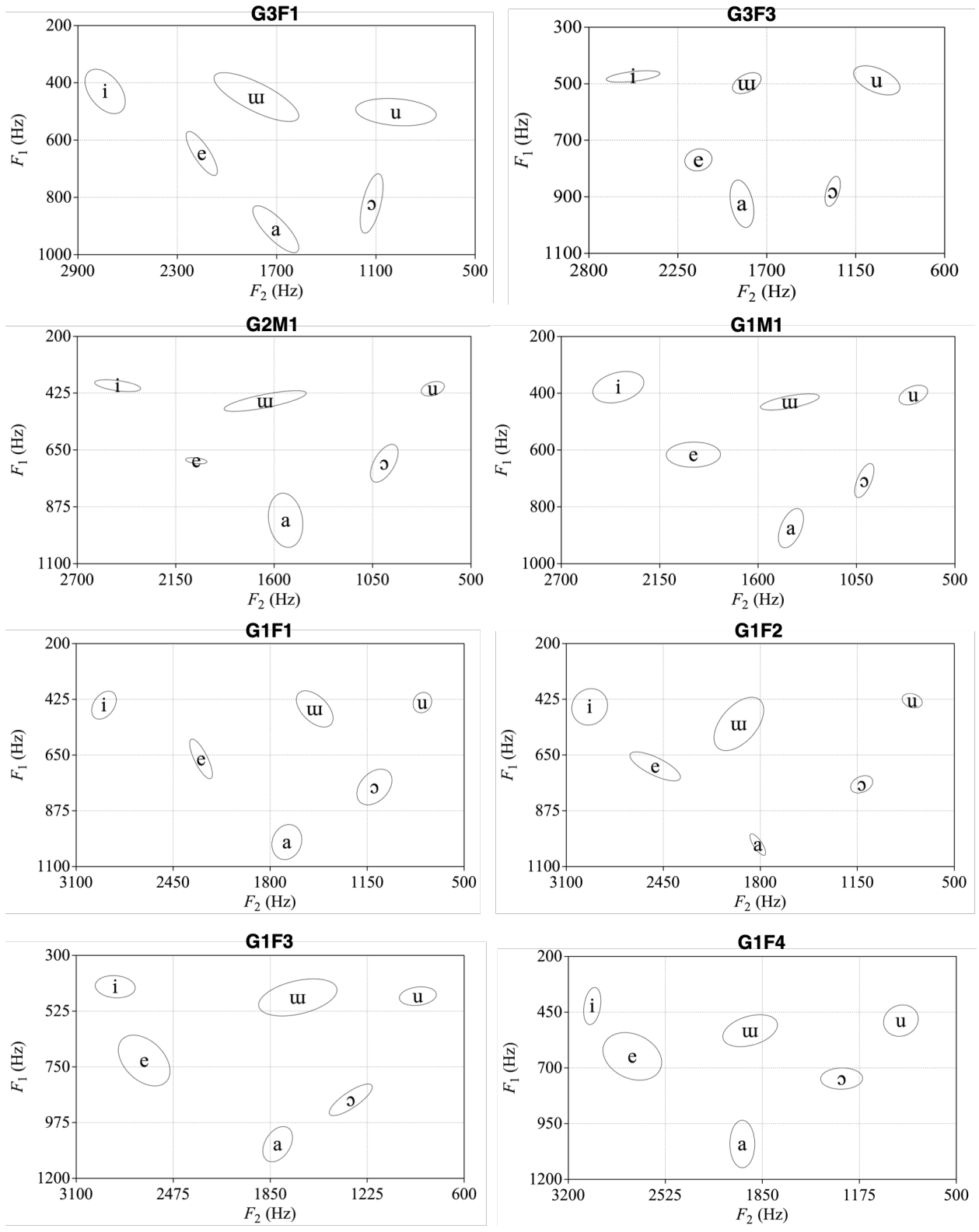


Figure 4.3: Graphs showing vowel spaces ( $F_1 \times F_2$ ) for eight primary speakers

While the cardinal *u* is a close, back, unrounded vowel, the graphs in Figure 4.3 show that acoustically, this vowel in CT is more of a close, central vowel. However, auditorily the vowel sounds more like *u* than the cardinal close, central, unrounded vowel *i* and therefore it will be transcribed as /u/ throughout this dissertation. There are two allophones to this phoneme. It is pronounced [u] in open syllables such as 魚 [hu<sup>55</sup>] ‘fish’ and it is pronounced [ə] in syllables ending in /ŋ/ such as 飯 [pəŋ<sup>11</sup>] ‘rice’. There were no tokens of this vowel in words ending in /p/, /k/, /ŋ/, or /m/ so the vowel’s behavior is unknown in those environments.

Allophones of /u/

/u/ → [u] / \_#

/u/ → [ə] / \_ŋ

A distinctive feature of Teochew phonology, including that of CT, is the presence of contrastive nasal vowels, which exist in monophthongs, diphthongs, and triphthongs. These vowels resulted from a sound change from nasal codas. Thus while 山 ‘mountain’ is *shān* in Mandarin, and *saan*<sup>1</sup> in Cantonese, it is /suã<sup>33</sup>/ in Teochew. This nasalization is phonemic. The contrast between oral and nasal monophthongs can be seen in the following minimal pairs.

/i/ and /ĩ/:

姨 /i<sup>55</sup>/ ‘aunt’

圓 /ĩ<sup>55</sup>/ ‘round’

/e/ and /ẽ/:

爬 /pe<sup>55</sup>/ ‘crawl’

平 /pẽ<sup>55</sup>/ ‘flat; same’

/a/ and /ã/:

膠 /ka<sup>33</sup>/ ‘plastic’

柑 /kã<sup>33</sup>/ ‘orange’

#### 4.2.1 Diphthongs

Cambodian Teochew has ten oral diphthongs. Nine of them have nasal counterparts.

Oral and nasal diphthongs and example words are showcased in Table 4.3.

Oral Diphthong	Example	Nasal Diphthong	Example
ai	菜 tɕ <sup>h</sup> ai <sup>21</sup> ‘vegetables’	aĩ	千 tɕ <sup>h</sup> ai <sup>33</sup> ‘thousand’
au	狗 kau <sup>52</sup> ‘dog’	aũ	好 hau <sup>21</sup> ‘willing’
ia	企 k <sup>h</sup> ia <sup>24</sup> ‘stand’	iã	行 kiã <sup>55</sup> ‘to walk’
iɔ	桥 kiɔ <sup>55</sup> ‘bridge’	iõ	薑 kiõ <sup>33</sup> ‘ginger’
iu	油 iu <sup>55</sup> ‘oil’	iũ	休息 hiũ <sup>33</sup> sek <sup>2</sup> ‘to rest’
ɔi	底 tɔi <sup>52</sup> ‘inside’	õĩ	看 t <sup>h</sup> õĩ <sup>52</sup> ‘to see’
ɔu	路 lɔu <sup>11</sup> road	X	
ua	沙 sua <sup>33</sup> ‘sand’	uã	山 suã <sup>33</sup> ‘mountain’
uɛ	粿 kue <sup>52</sup> ‘rice cake’	uẽ	苹果 p <sup>h</sup> ɛŋ <sup>11</sup> kuẽ <sup>52</sup> ‘apple’
ui	贵 kui <sup>21</sup> ‘expensive’	uĩ	危 kuĩ <sup>55</sup> ‘tall’

Table 4.3: Diphthongs in Cambodian Teochew

Khoo (2017c) found that some words in the five studied Malaysian Teochew varieties were pronounced without nasalization, though they were historically known to be nasalized. The same phenomenon was seen in words in the surveyed speakers of CT. As Khoo showed, there is no predictable pattern as to this reduced nasalization. Thus while a character like 爱 /aĩ<sup>52</sup>/ ‘to love’ was nasalized in all five Malaysian varieties, the same diphthong was only nasalized in three varieties in the character 指 [tsaĩ<sup>52</sup>] ~ [tsai<sup>52</sup>] ‘finger’ (Khoo, 2017c, p. 22). Khoo stated that the second character in the word for ‘tiger’ 老虎 was nasalized in all five studied varieties. However, none of the tokens in this study (from G1M1, G1F1, G1F4) were said with nasalization, with speakers instead pronouncing it /lau<sup>24-11</sup>hɔu<sup>52</sup>/. No other words were

found with an *ɔ̃* diphthong and therefore it is not included in this phonology of CT. While we included the *aĩ* diphthong, there was only one word found, 好 *haĩ*<sup>21</sup> ‘willing’.

#### 4.2.3 Triphthongs

Cambodian Teochew has two oral triphthongs. They each have a nasalized counterpart, though very few attested words have been found for each nasal triphthong. Triphthongs in CT are presented in Table 4.4.

Oral Triphthong	Example	Nasal Triphthong	Example
iau	鳥 tɕiau <sup>52</sup> ‘bird’	iãu	𠵼 iãu <sup>24</sup> ‘lift, open’
uai	歪 uai <sup>33</sup> ‘slanted’	uãi	樣 suãi <sup>11</sup> ‘mango’

Table 4.4: Triphthongs in Cambodian Teochew

### 4.3 Tone

Cambodian Teochew is a tonal language and all syllables have a tone. When a word or syllable is spoken in isolation, its tone is called the dictionary/citation tone. CT has eight citation tones which are covered in §4.3.1. A syllable with the same meaning can have a different tone depending on the construction it is in. This phenomenon, known as tone sandhi, is explored in §4.3.2.

#### 4.3.1 Citation tone

CT’s eight citation tones are presented in Table 4.5 with example words.

Tone Name	Pitch Description		Example Words	
yin ping 阴平	33	mid level	东 tɑŋ <sup>33</sup> ‘east’	刀 tɔ <sup>33</sup> ‘knife’
yang ping 阳平	55	high level	虫 t <sup>h</sup> ɑŋ <sup>55</sup> ‘worm’	桃 tɔ <sup>55</sup> ‘type of fruit’
yin shang 阴上	52	high falling	等 tɑŋ <sup>52</sup> ‘wait’	短 tɔ <sup>52</sup> ‘short’
yang shang 阳上	24	low rising	重 tɑŋ <sup>24</sup> ‘heavy’	弟 ti <sup>24</sup> ‘younger brother’
yin qu 阴去	21	low falling	冻 tɑŋ <sup>21</sup> ‘frozen’	倒 tɔ <sup>21</sup> ‘to pour’
yang qu 阳去	11	low level	慢 maŋ <sup>11</sup> ‘slow’	袋 tɔ <sup>11</sup> bag
yin ru 阴入	2	low checked	隻 tɕia <sup>?</sup> 2 ‘animal classifier’	室 sik <sup>2</sup> ‘room’
yang ru 阳入	5	high checked	食 tɕia <sup>?</sup> 5 ‘to eat’	翼 sik <sup>5</sup> ‘wing’

Table 4.5: The eight basic tones of Cambodian Teochew

Pitch description describes the behavior of the tones. Following Chao (1968), numerical superscripts represent the approximate pitch level of the voice on a five-point scale with [1] representing the low point and [5] the high point. Level tones more or less stay at the same approximate pitch level throughout the tone bearing unit while the pitch level decreases in the falling tones and increases in the rising tone. The exact pitch level on a five-point scale may vary per speaker but the values in Table 4.5 were chosen as benchmarks for CT, based on the nine primary speakers. Overall, the behavior of the tones will match. For example, while a speaker might have pronounced a word like 九 [kau<sup>53</sup>] ‘nine’ in close transcription, it will be transcribed as /kau<sup>52</sup>/ because [52] represents the phoneme for the yin shang or high falling tone.

While there are 8 unique tones, the yin ru and yang ru tones only appear on non-nasal closed syllables. These tones are shorter in length which is represented by the singular digit indicating the tone. The following word sets show the contrasting environments of the checked and level tones.

翼 /sik <sup>5</sup> / ‘wing’	骨 /kuk <sup>2</sup> / ‘bone’
时 /si <sup>55</sup> / ‘time, period’	旧 /ku <sup>11</sup> / ‘old’
丝 /si <sup>33</sup> / ‘silk’	

Pitch values and ranges in Hertz (Hz) for each tone are not absolute as they will be different for each speaker. Plotting smoothed pitch curves per speaker on a pitch (Hz) x time graph is a good way to visualize the relation between the different tones in a language. Figure 4.4 showcases graphs plotting the behavior of the citation tones for eight primary speakers. Using picture stimuli and elicitation data, tones from words said in isolation were analyzed using Praat. Tone bearing units were manually labeled in an interval tier of a TextGrid. Minimally five tokens per tone were selected. The analyze\_tone script from Ring (2017) was used to extract the pitch values at different points in the syllable, by percentage increments. The highest number of increments was chosen such that the script ran successfully. There were five (0, 20, 40, 60, 80) points for speaker G2M1, seven (0, 14, 29, 43, 57, 71, 86) for G1M1 and G1F1, nine (0, 11, 22, 33, 44, 55, 66, 77, 88) for G3F1, and ten (0, 10, 20, 30, 40, 50, 60, 70, 80, 90) for G3F3, G1F2, G1F3, and G1F4. The same Ring (2017) script subsequently took that data, normalized it, and plotted it on a graph of pitch (Hz) x time (seconds). The results of the scripts are presented in Figure 4.4.

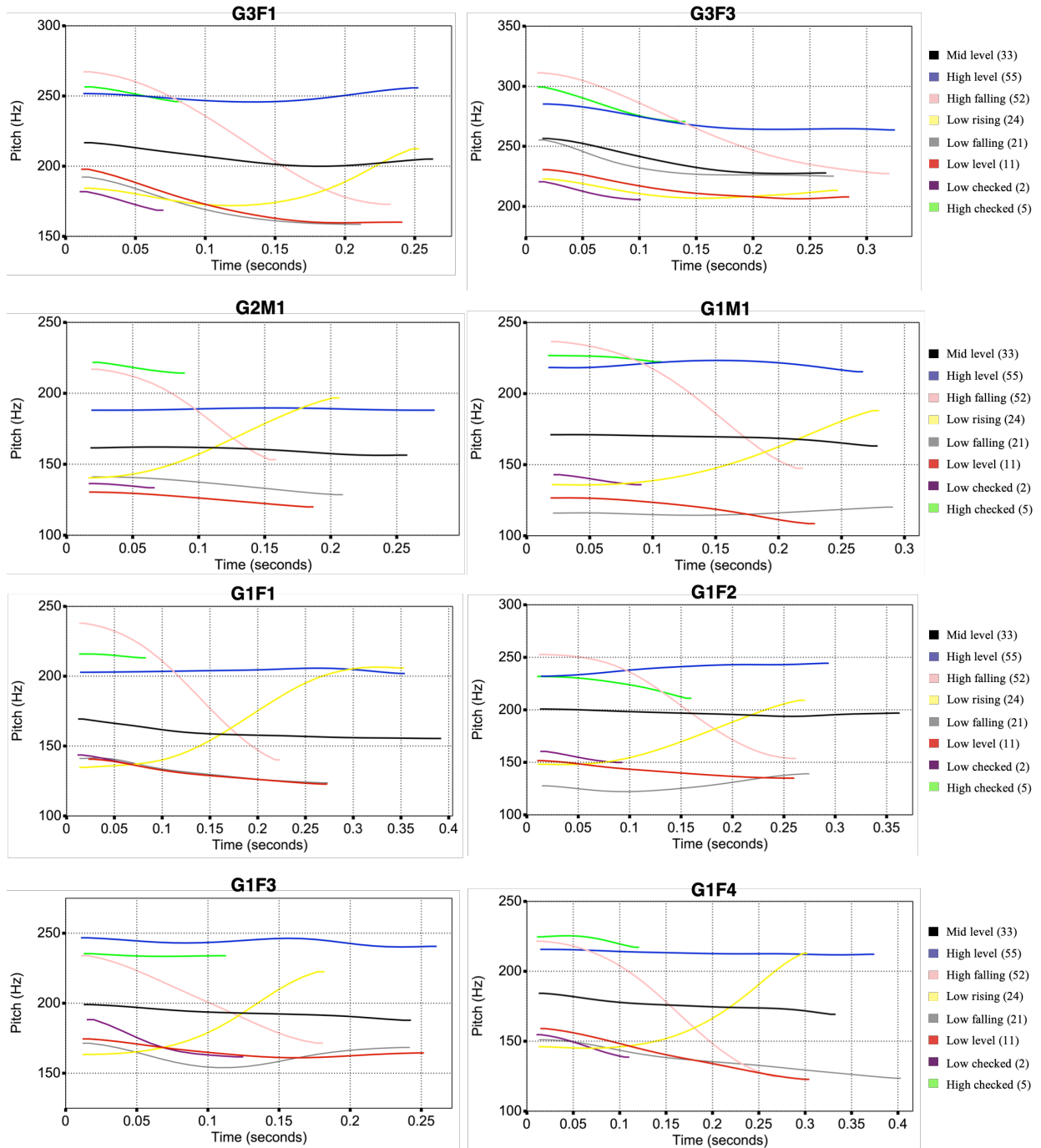


Figure 4.4: Graphs showing citation tone pitch (Hz) x time (seconds) for eight primary speakers

The graphs confirm that the checked tones (green and purple) are of shorter duration than the non-checked tones. Additionally, the variation in pitches between speakers is clear. For example, while the pitch of G1M1's high level tone is around 225 Hz, a tone at 225 Hz would

be a mid level tone for speaker G3F1, and a low level tone for speaker G3F3. This is why the pitches of the tones relative to each other for any one speaker is what is important, rather than how they compare to another speaker's pitches.

In Figure 4.4, the yin qu or low falling tone (grey) is very similar to the yang qu low level tone (red) for nearly all of the speakers. Historically, words pronounced in this tone had a dipping pitch or contour tone. There is evidence that these words are still pronounced with a dipping pitch in Chaozhou, China, as shown in the graph in Figure 4.5 from Hong (2013).

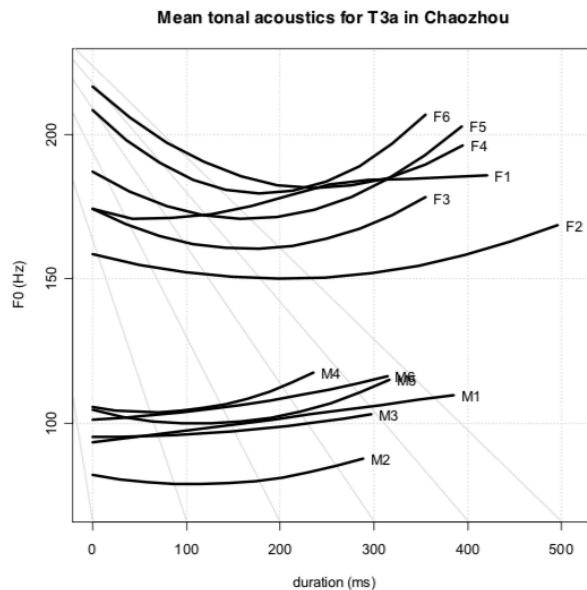


Figure 4.5: Pitch (Hz) x duration (ms) graph of the yin qu tone from twelve speakers in Hong's (2013, p. 241) study

Speakers G1M1, G1F2, and G1F3 do appear to have a slight rise at the end of this tone, though the other five speakers do not. Despite the tone's acoustic similarity to the low level tone for most of the CT speakers in this study, it is still viewed by speakers as a unique tone and it behaves differently in sandhi environments as will be shown in the next section.

Therefore, I am describing this tone as low falling and words with this tone will be transcribed with [21].

#### 4.3.2 Tone sandhi

CT has an extensive system of tone sandhi whereby the tone of a syllable may be pronounced differently whether in isolation or in another word or phrase. Tone sandhi is represented per convention where a syllable's citation tone appears first in the superscript, followed by a hyphen, and then the syllable's surface tone after undergoing tone sandhi. For example, when 牛 /gu<sup>55</sup>/ 'cow' is spoken in isolation in a reference construction meaning 'cow', it is pronounced with a high level tone. When the same word is used in a modification construction, such as 牛肉 /gu<sup>55-11</sup>ba<sup>2</sup>/ 'cow-meat or beef', the *gu* is pronounced with a low level tone. The general tone sandhi pattern is that in multi-syllable words, the final syllable will be pronounced with the citation tone while the preceding syllables are pronounced with sandhi tones<sup>26</sup>. Table 4.6 shows the tone sandhi patterns in CT with examples. The final column illustrates that when the same syllable is reduplicated, the initial syllable is pronounced in sandhi tone.<sup>27</sup>

---

<sup>26</sup> More nuanced tone sandhi rules likely exist in Cambodian Teochew. However, since detailed phonetic analysis was not the aim of this grammar, they will not be covered here. For more comprehensive information and acoustic analyses of tone sandhi patterns in Southern Min, see Lin (2019).

<sup>27</sup> This is the expected behavior in reduplicated syllables, though it does not always occur for some CT speakers.

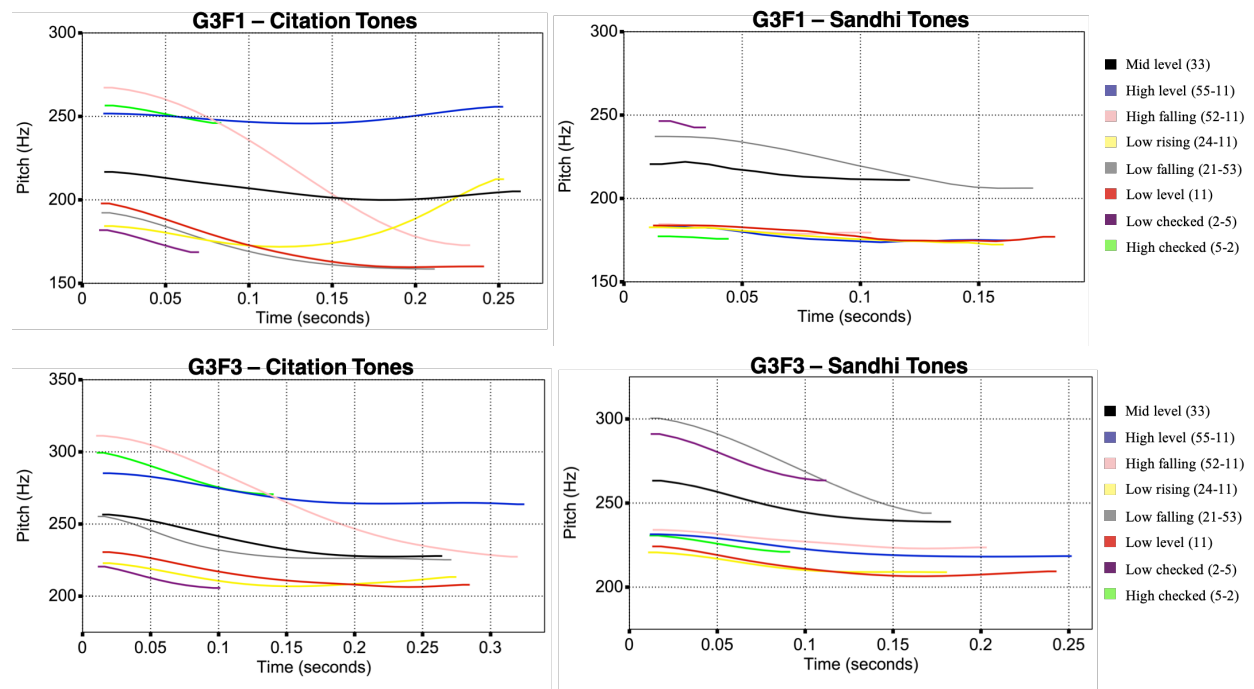
Citation Tone	Example	Sandhi Tone	Example in Sandhi Environment	Reduplication
33	乌 ɔu <sup>33</sup> ‘black’	33	乌色 ɔu <sup>33</sup> sek <sup>2</sup> ‘black’	真真 tɕiŋ <sup>33</sup> tɕiŋ <sup>33</sup> ‘extremely’
55	红 aŋ <sup>55</sup> ‘red’	11	红色 aŋ <sup>55-11</sup> sek <sup>2</sup> ‘red’	圆圆 ĩ <sup>55-11</sup> ĩ <sup>55</sup> ‘round circle’
52	底 tɔi <sup>52</sup> ‘inside’ 海 hai <sup>52</sup> ‘sea’	24	底畔 tɔi <sup>52-24</sup> paĩ <sup>55</sup> ‘inside’ 海水 hai <sup>52-24</sup> tɕui <sup>52</sup> ‘seawater’	久久 ku <sup>52-24</sup> ku <sup>52</sup> ‘a very long time’
		11	底畔 tɔi <sup>52-11</sup> paĩ <sup>55</sup> ‘inside’ 海舷 hai <sup>52-11</sup> ki <sup>55</sup> ‘beach’	好好 ho <sup>52-11</sup> ho <sup>52</sup> ‘very good’
24	雨 hɔu <sup>24</sup> ‘rain’ 十二 tɕap <sup>5-2</sup> dzi <sup>24</sup> ‘twelve’	11	雨遮 hɔu <sup>24-11</sup> tɕia <sup>33</sup> ‘umbrella’ 二十 dzi <sup>24-11</sup> tɕap <sup>5</sup> ‘twenty’	弟弟 ti <sup>24-11</sup> ti <sup>24</sup> ‘younger brother’
21	嘴 tɕ <sup>h</sup> ui <sup>21</sup> ‘mouth’	53	嘴唇 tɕ <sup>h</sup> ui <sup>21-53</sup> tun <sup>55</sup> ‘lips’	细细 sɔi <sup>21-53</sup> sɔi <sup>11</sup> ‘very small’ 笑笑 tɕ <sup>h</sup> io <sup>21-53</sup> tɕ <sup>h</sup> io <sup>21</sup> ‘to smile’
11	大 tua <sup>11</sup> ‘big’	11	大石 tua <sup>11</sup> tɕioŋ <sup>5</sup> ‘boulder’	慢慢 maŋ <sup>11</sup> maŋ <sup>11</sup> ‘slowly’ 恬恬 tiam <sup>11</sup> tiam <sup>11</sup> ‘repeatedly’
2	蜀百 tɕek <sup>5-2</sup> peŋ <sup>2</sup> ‘one hundred’ 石 tɕioŋ <sup>5</sup> ‘stone’	5	蜀百一 tɕek <sup>5-2</sup> peŋ <sup>2-5</sup> ik <sup>2</sup> ‘one hundred one’ 石头 tɕioŋ <sup>5-2</sup> t <sup>h</sup> au <sup>55</sup> ‘stone’	出出 tɕ <sup>h</sup> uŋ <sup>2-5</sup> tɕ <sup>h</sup> uŋ <sup>2</sup> ‘to go out’
5	食 tɕiaŋ <sup>5</sup> ‘to eat’ 白 peŋ <sup>5</sup> ‘white’	2	食饭 tɕiaŋ <sup>5-2</sup> puŋ <sup>11</sup> ‘to eat’ 白色 peŋ <sup>5-2</sup> sek <sup>2</sup> ‘white’	日日 dzik <sup>5-2</sup> dzik <sup>5</sup> ‘everyday’

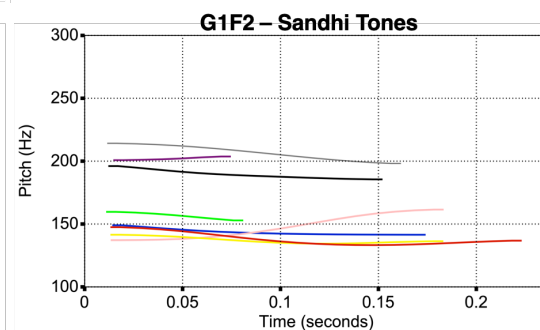
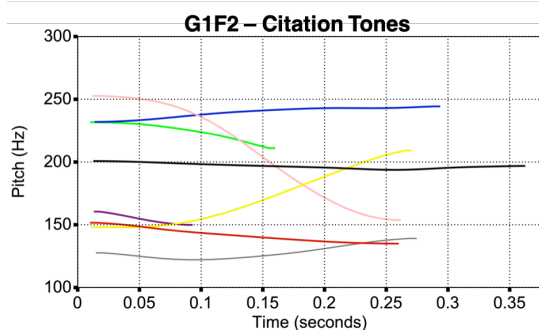
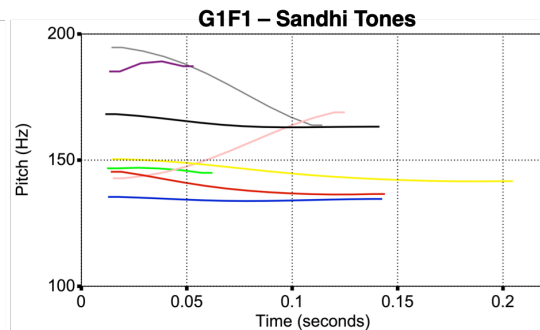
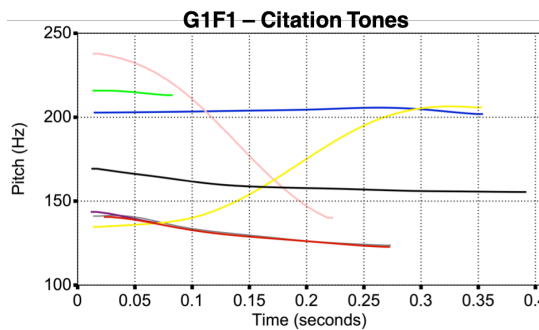
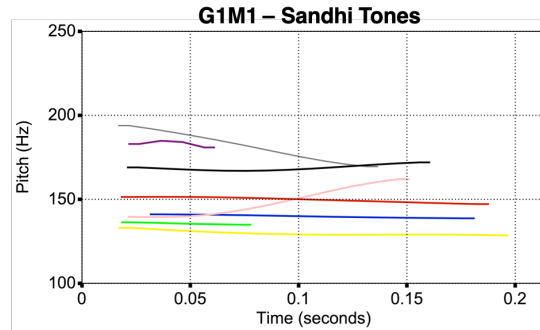
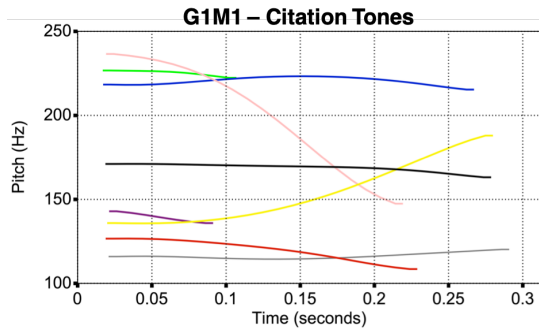
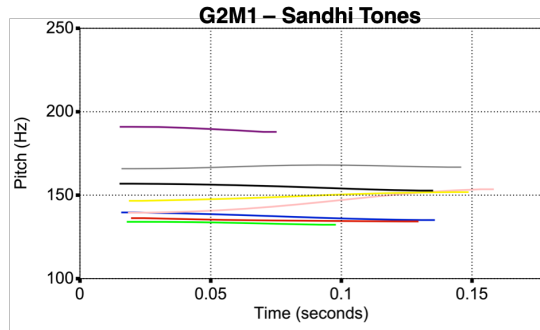
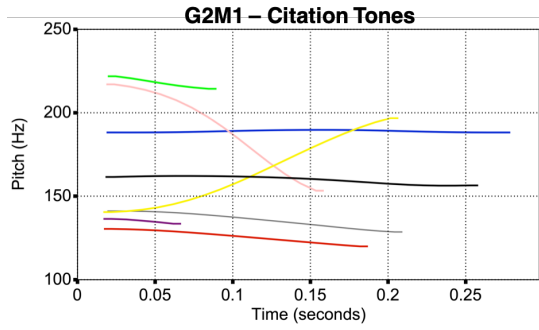
Table 4.6: Tone sandhi in Cambodian Teochew

Not all tones undergo tone sandhi. The mid and low level tones are pronounced with the same approximate relative pitch values in sandhi environments. Additionally, variation exists amongst the speakers consulted in this project as to the pronunciation of words with the yin shang high falling citation tone when in sandhi domains. The G1 speakers have a more clear

rise in this sandhi tone as in 顶畔 *teŋ*<sup>52-24</sup>*paŋ*<sup>55</sup> ‘atop’. For the G3 speakers, this tone has merged with the low level tone in sandhi environment as in 顶畔 *teŋ*<sup>52-11</sup>*paŋ*<sup>55</sup> ‘atop’. This variation will be noted throughout the dissertation. Since there are only three G3 speakers in this study, it would be interesting to see if there is a broader generational pattern in that tone’s behavior.

A similar procedure to the one outlined in §4.3.1 was followed to plot the sandhi tones. Data from picture stimuli/elicitation sessions and a selected text for each speaker was analyzed using Praat. Tone bearing units in sandhi domains were manually labeled in an interval tier of a TextGrid. The analyze\_tone script from Ring (2017) was used to extract the pitch values at five (G3F1, G2M1, G1M1, G1F3), six (G1F1), seven (G1F2), nine (G3F3), or ten (G1F4) points in the syllable. The same script then normalized the data and plotted it on a graph of pitch (Hz) x time (seconds). Figure 4.6 shows a side by side comparison of citation and sandhi tone graphs for eight speakers.





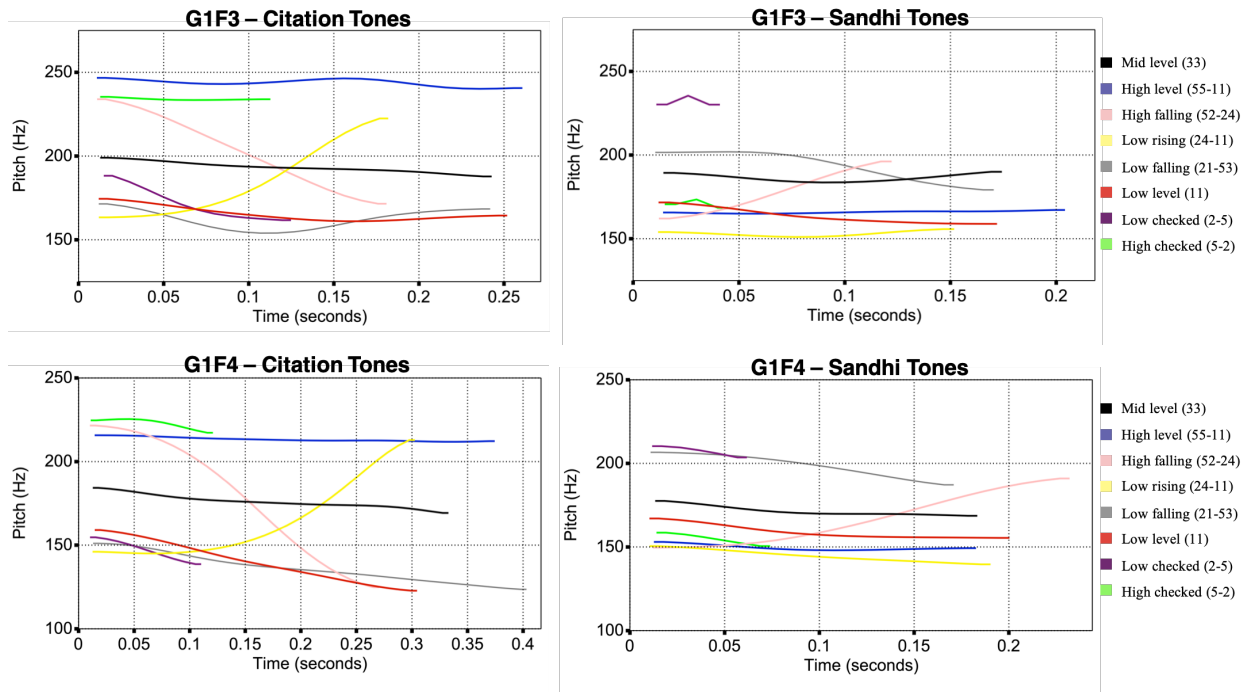


Figure 4.6: Graphs showing citation and sandhi tone comparisons for eight speakers

Overall, sandhi syllables tend to be shorter, as evidenced by the reduced time on the x axes in the graphs on the right. The difference in the sandhi pronunciation of the high falling tone (pink) between the G1 and G3 speakers can also be seen in Figure 4.6.<sup>28</sup>

#### 4.4 Syllable Structure

The syllable structure for Cambodian Teochew is illustrated in Figure 4.7.

<sup>28</sup> While there was not sufficient data suitable to producing tone graphs for speaker G3F2, auditorily her tones behave similarly to the other G3 speakers.

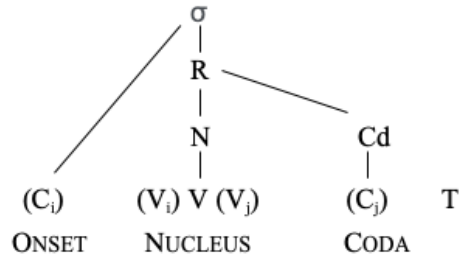


Figure 4.7: Syllable structure in Cambodian Teochew

Where:

$C_i$  = [-syllabic], any consonant

$V_i$  = [-syllabic, +high, +front/+back], /i/ or /u/

$V$  = [+syllabic], any vowel or syllabic nasal

$V_j$  = [-syllabic, +high, +front/+back], /i/ or /u/

$C_j$  = [-voice, -delayed release, -coronal]/[+nasal, -coronal], /p/, /ʔ/, /k/, /m/ or /ŋ/

$T$  = any tone

A syllable minimally contains a vowel or syllabic nasal, and a tone. Diphthongs and triphthongs can exist with the high vowels /i/ or /u/. Coda is restricted to /p/, /ʔ/, /k/, /m/ or /ŋ/. Rhyme is restricted to three segments, so there cannot be a consonant,  $C_j$ , after a triphthong. Consonant clusters are not found. Table 4.7 gives the possible syllable structures in Cambodian Teochew with example words.  $VV_jC_j$  and  $C_iVV_jC_j$  do not appear very productive, with only one word found for each structure.

Structure	Examples
V	𣦵 ĩ <sup>24</sup> ‘to sleep’ 唔 m <sup>11</sup> ‘NEG’
V <sub>i</sub> V	贏 iã <sup>55</sup> ‘to win’ 碗 uã <sup>52</sup> ‘bowl’
VV <sub>j</sub>	鞋 ɔi <sup>55</sup> ‘shoe’ 乌 ɔu <sup>33</sup> ‘black’
V <sub>i</sub> VV <sub>j</sub>	歪 uai <sup>33</sup> ‘slanted’ 𠵿 iãu <sup>24</sup> ‘lift, open’
VC <sub>j</sub>	鴨 aɿ <sup>2</sup> ‘duck’ 红 aŋ <sup>55</sup> ‘red’
V <sub>i</sub> VC <sub>j</sub>	药 ioɿ <sup>5</sup> ‘medicine’ 盐 iam <sup>55</sup> ‘salt’
VV <sub>j</sub> C <sub>j</sub>	狭 ɔiɿ <sup>5</sup> ‘narrow’
C <sub>i</sub> V	马 be <sup>52</sup> ‘horse’ 耳 hĩ <sup>24</sup> ‘ear’
C <sub>i</sub> V <sub>i</sub> V	车 tɕ <sup>h</sup> ia <sup>33</sup> ‘car’ 瓜 kue <sup>33</sup> ‘melon’
C <sub>i</sub> VV <sub>j</sub>	看 t <sup>h</sup> ɔĩ <sup>52</sup> ‘to see’ 头 t <sup>h</sup> au <sup>55</sup> ‘head’
C <sub>i</sub> V <sub>i</sub> VC <sub>j</sub>	热 juaɿ <sup>5</sup> ‘hot’ 凉 liaŋ <sup>55</sup> ‘cool’
C <sub>i</sub> VV <sub>j</sub> C <sub>j</sub>	八 pɔiɿ <sup>2</sup> ‘eight’
C <sub>i</sub> VC <sub>j</sub>	骨 kuk <sup>2</sup> ‘bone’ 重 taŋ <sup>24</sup> ‘heavy’
C <sub>i</sub> V <sub>i</sub> VV <sub>k</sub>	猫 ŋiau <sup>33</sup> ‘cat’ 𣞤 suãi <sup>11</sup> ‘mango’

Table 4.7: Syllable structure examples in Cambodian Teochew

#### 4.5 Summary

This chapter presented an approximate phonology of Cambodian Teochew based on the speakers consulted in this study. Cambodian Teochew has 18 consonant phonemes, six oral and three nasal monophthongs, ten oral and nine nasal diphthongs, and two oral and nasal triphthongs. There are eight basic tones: high, mid, and low level, high falling, low falling, low rising, and low and high checked tones. The two checked tones are in complementary distribution with the low and high level tones. All tones except the low and mid-level tones undergo tone sandhi when they are not the final syllable in a multi-syllable word or phrase. Syllables in Cambodian Teochew are minimally a vowel or syllabic nasal, and a tone. Diphthongs and triphthongs must contain an /i/ or /u/. Coda is restricted to non-coronal voiceless unaspirated or nasal stops.

When comparing the presented CT phonology to the literature on other Teochew varieties around the region, a few differences are found. As for consonants, the alveolo-palatal affricates were not found to be phonemic in the other varieties, and instead were noted as allophonic variation of alveolar affricates before front vowels. As first posited in McFarland (2017), the sound change from alveolar to alveolo-palatal affricates in Cambodian Teochew could be a natural internal sound change resulting from spreading of the allophonic palatalization before front vowels to the rest of the vowels. Alternatively, it could be a result of influence from Khmer, which has several palatal phonemes including /c/, /ɲ/, /j/ (Huffman, 1967). Gruzdeva (2015, pp. 159-160) noted an increase in palatal consonants in Nivkh due to influence from Russian, the majority language.

X. Chen noted the existence of an initial /f/ phoneme in borrowed transliterated words in the Teochew spoken in Johor Bahru, Malaysia (2003, p. 43), in Ho Chi Min, Vietnam and Vientiane, Laos (2009, 2014), in Pontianak, Indonesia (2004, p. 61), and in Chiang Mai and

Hat Yai, Thailand (2019, p. 17). Examples of such words include  $\sigma^{33}f\theta^{53}$  ‘offer’ (X. Chen, 2003, p. 44),  $fak^5$  ‘fax’ (X. Chen, 2009, p. 17), and  $\sigma^{33}f^{33}si^{11}$  ‘office’ (X. Chen, 2014, p. 61). Chen did not find this sound in the Teochew spoken in Cambodia (2009), or in Bangkok Teochew of Thailand (2014, 2019). In Goh (2020), there are no words or loanwords with /f/ in Singapore Teochew. Instead, words use other strategies to achieve that pronunciation such as in 羔丕  $ko^{33}pi^{55}$  ‘coffee’, 菲林  $hui^{33}lim^{55}$  ‘film’, 乌必  $\sigma^{33}p^{h}ik^5$  ‘office’, and 沙发  $sa^{33}huak^l$  ‘sofa’.

Conversely, X. Chen (2014, p. 57) does include words such as  $fom^{33}$  ‘form’,  $f\alpha^{33}$  ‘fax’, and  $fui^5po^{11}$  ‘football’ in Singapore Teochew. There were some instances of speakers in this study using words with *f*. Speaker G1F1 used the Mandarin word 非洲  $f\bar{e}i\ zh\bar{o}u$  ‘Africa’, but then corrected herself to the Teochew /hui<sup>33</sup>tɕiu<sup>33</sup>/. G3F1 said *filiks* ‘Felix’ throughout her telling of the *Looking for Felix* story. This is more likely an example of code-switching as the word violates CT phonology since consonant clusters are not allowed. While Huffman (1967) lists /f/ as a phoneme in borrowed words in Khmer, the sound is not used at all in Filippi and Hiep’s (2016) Khmer pronouncing dictionary. Therefore, since Khmer does not have a high frequency of *f* sounds, borrowed Khmer words would not introduce that sound to the CT phonology. However, as the younger speakers in Cambodia have more exposure to English and Mandarin, it would be unsurprising if the *f* sound starts appearing more frequently in loanwords in Cambodian Teochew.<sup>29</sup>

As for vowels, there is a lot of variation in the literature as to the pronunciation of the vowel in a word like 鱼 /hu<sup>55</sup>/ fish. In X. Chen (2003, 2009, 2014, 2019), this vowel is /ɤ/. Hong (2013) calls it /ə/ based on her acoustic analyses. In Singapore Teochew, Yeo (2011) says the vowel is /ə/ while Low (2014) has /i/. In some varieties, this vowel has also been nasalized.

---

<sup>29</sup> However, Cambodian Teochew seems resistant to lexical change (see McFarland, 2021).

Thus, Singapore Teochew has words like 黃 /ə̃<sup>55</sup>/ ‘yellow’ and 遠 /hə̃<sup>35</sup>/ ‘far’ (Goh, 2020; Yeo, 2011), as compared to CT’s [əŋ<sup>55</sup>] ‘yellow’ and [həŋ<sup>24</sup>] ‘far’. The nasalized coda has been absorbed into the vowel in ST, but that sound change has not (yet) occurred in Cambodian Teochew. X. Chen (2003) notes the Teochew of Johor Bahru includes some words with *õ* as in 短 *tõ<sup>53</sup>* ‘short’, 狀 *tsõ<sup>35</sup>* ‘shape’, 二 *lõ<sup>35</sup>* ‘two’, and 秆 *kõ<sup>53</sup>* ‘stalk’. X. Chen also included a nasalized *õ* final in the Vietnamese, Laotian, and Cambodian varieties (2009), as well as in Singapore and Pontianak Teochew (2014), and in the Bangkok, Chaing Mai, and Hat Yai varieties of Thailand (2019). However, in this study, there was no evidence that the /o/ vowel was nasalized in monophthongs in Cambodian Teochew. Cambodian Teochew also has two fewer nasal diphthongs than some of the other varieties. And while there are examples of words with nasal triphthongs, they do not seem to be very productive.

Finally, the CT tones were slightly different than some other varieties. The historically contour tone has been reduced to a low falling tone for six of nine speakers. This tone has also lost its contour for some speakers in Singapore Teochew. Yeo (2011) transcribes it as [13] for the speakers in her study, but Low (2014) and Goh (2020) maintain it as [213]. A brief analysis of a Singapore Teochew text (Learn Dialect in Singapore, 2020) using the Ring (2017) script found the tone in a word like 去 /ku<sup>21</sup>/ ‘to go’ in phrase final position to be very similar to the low level tone. Syllables with this tone still behave differently in sandhi environments in both Cambodian Teochew and Singapore Teochew. One difference from other varieties in sandhi tones was found in the behavior of the high falling tone. It has merged with the low level tone in sandhi environments for three of nine speakers. Since it was the G3 speakers who spoke this way, this could be a generational change. If the younger speakers have less exposure to the

language, there may be fewer opportunities to learn the nuanced rules of tone sandhi. More research is needed to confirm this hypothesis.

## Chapter 5 – Reference and Referent Expressions

Recall from Chapter 1 the three propositional act functions of reference, modification, and predication. This chapter covers referring phrase constructions which perform the act of reference, where reference is what the speaker is talking about (Croft, forthcoming, p. 33).

Referring phrase constructions can be object reference (§5.1), property reference (§5.4), or action reference (§5.5), though object reference is the prototypical referring expression. Croft (forthcoming) uses the term “noun” to mean reference to an object. The term is not used as a word class, but instead is a comparative concept (Haspelmath, 2010). The usage of the word “noun” cannot be extended beyond the function just described. Thus there is a distinction between the word *doctor* in *The doctor just walked in* and *She is a doctor*. The same word has a different function depending on the construction it appears in. This again is contrary to the word class approach where part of speech is inherent to the word, regardless of its position in a sentence. In RCG, *doctor* in *The doctor just walked in* is object reference (or “noun” according to Croft) while *doctor* in *She is a doctor* is object predication. As outlined in Chapter 1 of this dissertation, using traditional terms like noun, verb, adjective, etc., even when defined clearly as above, can still be problematic and confusing because it can be hard for readers to disassociate from the old usages of the terms. Therefore, Croft’s (forthcoming) terminology has been updated throughout this chapter to more accurately reflect the functional usages of each construction.<sup>30</sup>

---

<sup>30</sup> Since the functional terms reflect the combination of semantic content and information packaging, they must be multi-word terms, instead of the more convenient traditional one-word terms used by Croft (noun, pronoun, demonstrative, complement, etc.). This is unavoidable.

This chapter also looks at the information status continuum (§5.2) and covers different constructions in Cambodian Teochew based on a referent’s information status. I also explore the Extended Animacy Hierarchy (§5.3) and how it manifests in Cambodian Teochew.

## 5.1 CT Object Reference Constructions

Object reference, henceforth objR, represents the prototypical referring phrase constructions. Per Croft (forthcoming), objRs can be divided into three semantic categories, as outlined in Table 5.1.

Semantics of reference	Construction
category (type)	common objR (§5.1.1)
individual	proper objR (§5.1.2)
contextual	contextual objR <sup>31</sup> (§5.1.3)

Table 5.1: Semantics of reference to an individual (adapted from Croft, forthcoming, p. 59)

### 5.1.1 CT Common ObjR Constructions

Common objR constructions select an object based on its type or category. In (1), *ɲiau*<sup>33</sup> ‘cat’ is an example of a CT Common ObjR Construction.

- (1) 猫            𠵼  
       ɲiau<sup>33</sup>        ɿ<sup>24</sup>  
       cat            sleep  
       ‘A cat sleeps.’ [G1F1-9]

The speaker says *ɲiau*<sup>33</sup> ‘cat’ to refer to the cat-type object in the story and the hearer is able to recognize what object the speaker is referring to because they have an understanding of what objects would be cat-type.

---

<sup>31</sup> Contextual objRs are called “pronouns” in Croft (forthcoming).

CT Common ObjR Constructions can stand alone. When describing a picture, consultants would often point to items and name them. Each of the words in (2) is a CT Common ObjR Construction.

- (2) 鸟 .. 蛇 .. 月娘  
 tɕiau<sup>52</sup> .. tɕua<sup>55</sup> .. bueŋ<sup>5-2</sup>niŋ<sup>55</sup>  
 bird .. snake .. moon<sup>32</sup>  
 ‘Bird..snake..moon.’ [G1F1-10]

### 5.1.2 CT Proper ObjR Constructions

Proper objR constructions make reference to a specific individual, such as people, places, and entities we know (Croft, forthcoming, p. 59). CT Proper ObjR Constructions are illustrated in (3-5). The examples in (3) make reference to specific people, such as the speaker himself (3a) or characters in the story (3b-c). Places make up the CT Proper ObjR Constructions in (4). In (4a), *kim<sup>33</sup>t<sup>h</sup>aŋ<sup>2</sup>* refers to the capitol city of Cambodia, Phnom Penh. The individual syllables/characters of 金塔 *kim<sup>33</sup>t<sup>h</sup>aŋ<sup>2</sup>* mean 金 *kim<sup>33</sup>* ‘gold’ and 塔 *t<sup>h</sup>aŋ<sup>2</sup>* ‘pagoda’ or 金塔 ‘golden pagoda’ together. However, in this context, the combination of characters gives the CT Proper ObjR that makes reference to Phnom Penh. The sentences in (5) show reference to entities such as languages. In (5b), 国语 *kək<sup>2-5</sup>gu<sup>52</sup>* ‘Mandarin’ literally translates to country-language but again is a CT Proper ObjR in this context.

---

<sup>32</sup> Note that the pronunciation of 月娘 *bueŋ<sup>2</sup>niŋ<sup>55</sup>* ‘moon’ for this speaker (n = 3) differs from that of the other speakers who say *gueŋ<sup>2</sup>niŋ<sup>55</sup>* (G3F1, G3F3, G1F2, G1F3, G1F4, G1M1, G2M1). The g → b might be a sound change in process.

- (3) a. 我 名 X  
ua<sup>52</sup> miã<sup>55-11</sup> [name redacted]  
1SG be.named [name redacted]  
‘I am named [name redacted].’ [G2M1-5]
- b. 伊 名 金鱼  
i<sup>33</sup> miã<sup>55-11</sup> kim<sup>33</sup>hu<sup>55</sup>  
3SG be.named gold-fish  
‘He is named Gold Fish.’ [G2F1US-20]
- c. 觅 Felix  
tɕ<sup>h</sup>ue<sup>11</sup> Felik  
look.for Felix<sup>33</sup>  
‘Looking for Felix.’ [G3F1-2]
- (4) a. 我 来 ... 金塔 高棉  
ua<sup>52</sup> lai<sup>55-11</sup> ... kim<sup>33</sup>tʰa<sup>2</sup> kau<sup>33</sup>mian<sup>55</sup>  
1SG come ... Phnom Penh Cambodia (Khmer)  
‘I come to Phnom Penh, Cambodia.’ [G2M1-5]
- b. 伊人 父母 在 唐山 来  
ij<sup>33</sup> pe<sup>24-11</sup> bɔ<sup>52</sup> tɔ<sup>11</sup> ten<sup>33</sup>suã<sup>33</sup> lai<sup>55</sup>  
3PL father-mother LOC Tangshan come  
‘[My] parents, they came from Tangshan.’ [G1M2-24]
- (5) a. 潮州话 着 但  
tiɔ<sup>11</sup>tɕiu<sup>33</sup>ue<sup>11</sup> tiɔ<sup>75</sup> tã<sup>21</sup>  
Teochew must speak  
‘(We) must speak Teochew.’ [G1F6-24]
- b. ...在 但 国语  
...tɔ<sup>11</sup> tã<sup>21-53</sup> kɔk<sup>2-5</sup>gu<sup>52</sup>  
...PROG speak Mandarin  
‘...[you] are speaking Mandarin.’ [G1F6-24]

---

<sup>33</sup> ‘Felix’ was spoken in English but the construction would be the same with a local name.

### 5.1.3 CT Contextual ObjR Constructions

Contextual objR constructions reference individuals/things based on the context known to the speaker and the hearer. Examples of CT Contextual ObjR Constructions include CT Personal Contextual ObjR Constructions (§5.1.3.1) and CT Deictic Contextual ObjR Constructions (§5.1.3.2).

#### 5.1.3.1 CT Personal Contextual ObjR Constructions

Personal contextual objRs are defined by interlocutor role (speaker, addressee, other) in a speech act. A speaker would use first person contextual objRs to refer to themselves, second person contextual objRs to refer to the addressee/hearer, and third person contextual objRs to refer to others. CT Personal Contextual ObjR Constructions are given in Table 5.2. It is important to note that being a personal contextual objR is not an inherent property of the words in Table 5.2. Rather, the words become CT Personal Contextual ObjR Constructions according to the context they appear in, as shown in the examples (6-18) below. For example, these same words can be used in CT Object Modification Constructions, which are explored in the next chapter.

	Singular	Plural
First Person	我 ua <sup>52</sup>	俺 naŋ <sup>52</sup> 阮 uaŋ <sup>52</sup> (exclusive* <sup>34</sup> )
Second Person	汝 lu <sup>52</sup>	恁 niŋ <sup>52</sup>
Third Person	伊 i <sup>33</sup>	伊人 i <sup>33</sup> naŋ <sup>11</sup> / iŋ <sup>33</sup>

Table 5.2: CT Personal Contextual ObjR Constructions

---

<sup>34</sup> Seemingly only used by some speakers.

Due to the nature of the primary data gathered for this dissertation, (retelling of stories in the third person), there were very few tokens of First and Second Person Contextual ObjR Constructions. Some are given below and more are included from secondary sources.

The use of CT First Person Singular Contextual ObjR Constructions is showcased in (6-7) and CT Second Person Singular Contextual ObjR Constructions are shown in (7). The speaker (G2M1; 我 *ua*<sup>52</sup> ‘1SG’) is expressing his sentiment in (6). In the examples in (7), G2M1 (the speaker; 我 *ua*<sup>52</sup> ‘1SG’) is talking to G2M2US (the addressee; 汝 *lu*<sup>52</sup> ‘2SG’). Notice how the form of the CT Second Person Singular Contextual ObjR Construction, 汝 *lu*<sup>52</sup> ‘2SG’, remains the same in each of the utterances in (7), regardless of the Contextual ObjR’s position in the sentence. This can be seen with 我 *ua*<sup>52</sup> ‘1SG’ as well. In (7a), 我 *ua*<sup>52</sup> ‘1SG’ is performing the action while in (8) the actions are being performed on 我 *ua*<sup>52</sup> ‘1SG’, and the forms are identical.

(6) 我 欢喜  
ua<sup>52</sup> huã<sup>33</sup>hi<sup>52</sup>  
1SG happy  
 ‘I am happy.’ [G2M1-5]

(7) a. 我 邀 汝  
ua<sup>52</sup> kiau<sup>33</sup> lu<sup>52</sup>  
1SG guide 2SG  
 ‘I guide you [around Cambodia].’ [G2M1-5]

b. 我 问 汝 “欢喜 歪 来 高棉?”  
ua<sup>52</sup> muŋ<sup>11</sup> lu<sup>52</sup> “huã<sup>33</sup>hi<sup>52</sup> bɔi<sup>24</sup> lai<sup>55-11</sup> kau<sup>33</sup>mian<sup>55?</sup>”  
1SG ask 2SG happy NEG.can come Cambodia  
 ‘I ask you “Are [you] happy to come to Cambodia?”’ [G2M1-5]

c.     汝     咀     “欢喜”  
           luu<sup>52</sup>   tã<sup>21</sup>   “huã<sup>33</sup>hi<sup>52</sup>”  
           2SG   say   happy  
           ‘You say “[I am] happy”.’ [G2M1-5]

(8)   “𠵼            食 我. 放 我 落”  
       “mɔ<sup>11</sup>            tɕiaɪ<sup>ʔ5-2</sup> ua<sup>52</sup>.   paŋ<sup>21-52</sup> ua<sup>52</sup>   lɔ<sup>ʔ5</sup>”  
       NEG.can       eat   1SG.   put   1SG   down  
       “‘Don’t eat me! Put me down!’” [G2F1US-18]

An example sentence with a CT First Person Plural Contextual ObjR Construction and a CT Second Person Plural Contextual ObjR Construction is given in (9). For context, Speaker G1F6 is being interviewed for a news story and is addressing the audience using 恁 *niŋ*<sup>52</sup> ‘2PL’. She invites anyone who is interested to come and learn Teochew music at her school, saying they (her and her school; 俺 *naŋ*<sup>52</sup> ‘1PL’) will not charge a fee.

(9)   恁    有    底    吶            欲    来            学习,        爱    来  
       niŋ<sup>52</sup>   u<sup>25-11</sup>   ti<sup>11</sup>tiaŋ<sup>55</sup>        aɪ<sup>21-53</sup>   lai<sup>55-11</sup>        hak<sup>5-2</sup>sip<sup>5</sup>,   aɪ<sup>21-53</sup>   lai<sup>55-11</sup>  
       2PL   have   who            want   come        learn        want   come  
       练            只    个            潮州        曲    □<sup>35</sup>    ...  
       lian<sup>11</sup>        tɕi<sup>52</sup>   kai<sup>55-11</sup>        ti<sup>55-11</sup>tɕiu<sup>33</sup>   k<sup>h</sup>ek<sup>2</sup>   nɔ<sup>33</sup>    ...  
       practice    this   LW            Teochew   song   PRT    ...  
       俺    在    无    价    收  
       naŋ<sup>52</sup>   tɔ<sup>11</sup>   bɔ<sup>11</sup>   ke<sup>21-53</sup>   siu<sup>33</sup>  
       1PL   PROG   NEG   fee   collect  
       ‘[If] you (pl.) have anyone who wants to come learn, come practice this  
       Teochew song...we are not collecting a fee.’ [G1F6-24]

<sup>35</sup> I follow Bauer (2018) in using the “empty-box” as a means of indicating a morphosyllable for which no standard or historical character corresponding to the meaning of the words exists (p. 136).

In some languages, there are two first person plural contextual objR constructions. Inclusive (first person plural) contextual objR constructions include the speaker and the addressee while exclusive contextual objR constructions do not include the addressee and instead refer to the speaker and someone who is not the addressee. English does not have unique first person plural contextual objR constructions that reflect this semantic distinction, though the concept exists. The sentences in (10a) show an inclusive interpretation while those in (10b) show an exclusive interpretation (assuming the addressee is not also going to lunch with the speaker and their mother).

- (10) a. What are you doing now? We should go to lunch together.  
 b. My mother is coming to town tomorrow. We are going to lunch together.

Some evidence was found for the existence of a CT Exclusive First Person Plural Contextual ObjR Construction, 阮 *uaŋ*<sup>52</sup> 1PL-EXCL, at least for one speaker, while other evidence points to only one CT First Person Plural Contextual ObjR Construction for other speakers, though more data is needed to confirm this hypothesis.

Text #24, News Interview, contained a few CT First Person Plural Contextual ObjR Constructions. Speaker G1M2 was the only speaker who used 阮 *uaŋ*<sup>52</sup> ‘1PL-EXCL’, and it was used for exclusive contexts. There were three tokens of 阮 *uaŋ*<sup>52</sup> ‘1PL-EXCL’ being used by this speaker. For example, in (11), the speaker did not practice music with the addressee (either the interviewer or the audience), and instead practiced it with someone else.

- (11) 了 块 闲 时候 阮 就 来 练 乐 □  
 liau<sup>52</sup> kɔ<sup>21</sup> aĩ<sup>55</sup> si<sup>11</sup> hau<sup>11</sup> uaŋ<sup>52</sup> tɕiu<sup>11</sup> lai<sup>55-11</sup> liaŋ<sup>11</sup> gau<sup>5</sup> ma<sup>33</sup>  
 then CLF free.time time 1PL-EXCL then come practice music PRT  
 ‘When we had free time, then we came to practice music.’ [G1M2-24]

In contrast, the two other speakers in text #24 appear to use 俺 *naŋ*<sup>52</sup> ‘1PL’ for exclusive contexts. In (9) above from speaker G1F6, 俺 *naŋ*<sup>52</sup> ‘1PL’ can only be understood to be exclusive, excluding the addressee(s). The addressees are the people the speaker is inviting to come to the school (恁 *niŋ*<sup>52</sup> ‘2PL’). It is not expected that they would be the ones doing the collecting of the fee, therefore they are not included in the Contextual ObjR, 俺 *naŋ*<sup>52</sup> ‘1PL’.

In (12) from speaker G1F7, 俺 *naŋ*<sup>52</sup> ‘1PL’ appears in an Object Modification Construction where it is modifying 团 *t<sup>h</sup>uaŋ*<sup>55</sup> ‘troupe’. It is not likely that the addressees, neither the interviewer nor the audience, are a part of the drama troupe and thus would not be included in the Object 俺 *naŋ*<sup>52</sup> ‘1PL’. Although in this example the Object 俺 *naŋ*<sup>52</sup> ‘1PL’ is not an Object Reference Construction, given the parallels between the Contextual Objects in the two constructions (Object Reference and Object Modification shown in Chapter 6), it still should lend some evidence to the hypothesis presented here.

- (12) ...俺 个 潮 潮剧 团 来 只块 ...  
 ...naŋ<sup>52</sup> kai<sup>55-11</sup> tiə<sup>11</sup>- tiə<sup>11</sup>-kiaŋ<sup>55-2</sup> t<sup>h</sup>uaŋ<sup>55</sup> lai<sup>55-11</sup> tɕei<sup>52</sup>kə<sup>21</sup> ...  
 ...1PL LW TC- TC-drama troupe come here ...  
 ‘[I hope that] our Teochew drama troupe comes here.’ [G1F7-24]

In text #23, G1F5US is showing her daughter and the audience (as the video was intended for YouTube) around her garden. At the end, she addresses the audience saying that she has shown them around her garden followed by (13) below. The 俺 *naŋ*<sup>52</sup> ‘1PL’ does not include the addressee because the addressee is 汝 *luŋ*<sup>52</sup> ‘2SG’ in the same utterance.

- (13) 汝 合 俺 睇 着  
 lu<sup>52</sup> ka<sup>ʔ2-5</sup> naŋ<sup>52</sup> t<sup>h</sup>ɔi<sup>52-24</sup> tiɔ<sup>ʔ5</sup>  
 2SG with 1PL see RC  
 ‘You have seen [it] with us.’ [G1F5US-23]

There are several tokens from speaker G2F1US where she uses 俺 *naŋ*<sup>52</sup> ‘1PL’ in inclusive contexts. In (14), the mother is addressing her son and the 俺 *naŋ*<sup>52</sup> ‘1PL’ refers to both the mother and the son. Similarly in (15), the tiger is speaking to the fox and the 俺 *naŋ*<sup>52</sup> ‘1PL’ refers to both the tiger and the fox. Finally, in (16), the speaker (a fish) is addressing his friend (another fish) and suggesting they both go [to a new place to play].

- (14) 妈妈 合 个 细 但 “个 虫, 伊 无 平样 俺”  
 ma<sup>55-11</sup> ma<sup>55</sup> ka<sup>ʔ2</sup> kai<sup>55-11</sup> sɔi<sup>21</sup> tã<sup>21-53</sup> “kai<sup>55-11</sup> t<sup>h</sup>aŋ<sup>55</sup> i<sup>33</sup> bɔ<sup>55-11</sup> pẽ<sup>55-11</sup> i<sup>5</sup> naŋ<sup>52</sup>”  
 mama with IK small say IK worm 3SG NEG same 1PL  
 ‘Mama says to her small one, “The worm, he is not the same as us.”’ [G2F1US-17]

- (15) “我 无 欲 食 汝 了. 俺 来 物 朋友”  
 “ua<sup>52</sup> bɔ<sup>55-11</sup> ai<sup>21-53</sup> tɕia<sup>ʔ5-2</sup> lu<sup>52</sup> ɔu<sup>11</sup>. naŋ<sup>52</sup> lai<sup>55</sup> mue<sup>ʔ5-2</sup> peŋ<sup>55-11</sup> iu<sup>11</sup>”  
 1SG NEG want eat 2SG already. 1PL come make friend  
 “I don’t want to eat you anymore. We have come to be friends” [G2F1US-18]

- (16) “俺 去 无”  
 “naŋ<sup>52</sup> k<sup>h</sup>u<sup>21</sup> bɔ<sup>55</sup>”  
 1PL go NEG  
 “Let’s go, no?” [G2F1US-20]

The usages of 俺 *naŋ*<sup>52</sup> ‘1PL’ in (9) and (12-16) provide evidence that, at least for some speakers, 俺 *naŋ*<sup>52</sup> ‘1PL’ is used for both exclusive and inclusive contexts. Speakers interviewed for this project also report using 俺 *naŋ*<sup>52</sup> ‘1PL’ in both inclusive and exclusive contexts. The CT First Person Plural Exclusive Contextual ObjR Construction 阮 *uaŋ*<sup>52</sup> ‘1PL-EXCL’ might be

still in use by only some speakers, or it may be used only if someone really wants to emphasize exclusivity. More evidence is needed to determine the extent of the usage of this CT ObjR.

The inclusive and exclusive usage of 俺 *naŋ*<sup>52</sup> ‘1PL’ is contrary to the usages reported in other Teochew varieties. 俺 *naŋ*<sup>53</sup> ‘1PL’ is said to be used in inclusive contexts in Jieyang (Xu, 2007) and Singapore (Goh, 2020; Low, 2014; Yeo, 2011) Teochew. Meanwhile, the Exclusive First Person Plural Contextual ObjR, which varies in pronunciation, is reported as 阮 *uaŋ*<sup>53</sup> (Mogher); *ŋ*<sup>53</sup> (TCKnow LLC., 2015); *uaŋ*<sup>53</sup> in Jieyang (Xu, 2007, p. 74); and *ĩ*<sup>53</sup> (Low, 2014, p. 18), and *ẽ*<sup>53</sup> (Goh, 2020, p. 110; Yeo, 2011, p. 35) in Singapore Teochew. According to X. Chen (2003), the Johor Bahru Teochew variety does not distinguish between inclusive and exclusive First Person Contextual ObjR Constructions and instead solely uses 俺 *naŋ*<sup>53</sup> ‘1PL’ (p. 379).

The reduction in the number of CT Contextual ObjR Constructions, at least for some speakers, could be due to language induced change as there are not distinct inclusive/exclusive Khmer First Person Contextual ObjR Constructions. Alternatively, given that the same change is reported in Johor Bahru Teochew, it could be a simplification of the grammar due to reduced language exposure.

CT Third Person Contextual ObjR Constructions can be used to refer to a variety of ‘others’ outside of the speech act. These Constructions are not specific to gender. Thus 伊 *i*<sup>33</sup> ‘3SG’ can be used to refer to both female (17) and male (18) entities.

- (17) 老姐            □    耍内            伊            □    跔            耍内  
*lau*<sup>24-11</sup>*tɕe*<sup>52</sup>    *na*<sup>11</sup>    *suŋ*<sup>52-11</sup>*lai*<sup>24</sup>,    *i*<sup>33</sup>    *na*<sup>11</sup>    *k<sup>h</sup>u*<sup>11</sup>    *suŋ*<sup>52-11</sup>*lai*<sup>24</sup>  
 old-girl            PROG    play-house    3SG    PROG    bend.over    play-house  
 ‘The girl is playing house. She is bending over, playing house.’ [G3F1-1]

- (18) 只个 公 ... 伊 有 孙 来 喊 伊  
 tɕia<sup>53</sup> kəŋ<sup>33</sup> ... i<sup>33</sup> u<sup>24-11</sup> suŋ<sup>33</sup> lai<sup>55-11</sup> ham<sup>21-53</sup> i<sup>33</sup>  
 this.LW old.man ... 3SG have grandchild come beckon 3SG  
 ‘This old man... he has a grandchild who comes and beckons him.’ [G3F1-1]

CT Third Person Contextual ObjR Constructions can also be used to refer to animal entities. 伊 *i*<sup>33</sup> ‘3SG’ is shown in (19) and an example sentence with 伊人 *iŋ*<sup>33</sup> ‘3PL’ is given in (20). In (21), 伊人 *i<sup>33</sup>naŋ<sup>11</sup>* ‘3PL’ is used to refer to a group of six animals paddling a boat and then dancing. Of course, animals do not usually paddle boats, and a cat would not usually ‘ask’ a cow for help as in (20). In these instances, the animals have been personified for the story. Nonetheless, it is still acceptable for CT Third Person Contextual ObjR Constructions to refer to animals doing more traditional animal things, like the sentences in (19).

- (19) a. 只个 猫 伊 死 口 只 个 画  
 tɕia<sup>53</sup> ŋiau<sup>33</sup> i<sup>33</sup> ʔ<sup>24-11</sup> na<sup>11</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> ue<sup>11</sup>  
 this.LW cat 3SG sleep LOC this LW 3SG picture  
 ‘This cat, it sleeps in this picture.’ [G3F1-3]
- b. 猪 啊. 伊 食  
 tu<sup>55</sup> a<sup>33</sup>. 3SG eat  
 pig PRT. i<sup>33</sup> tɕia<sup>ʔ5</sup>  
 ‘[It’s] a pig. It eats.’ [G2M1-5]
- (20) 伊 罗 问 牛 互 牛 相辅. 伊人 过 溪  
 i<sup>33</sup> lo<sup>11</sup> muŋ<sup>11</sup> gu<sup>55</sup> həu<sup>11</sup> gu<sup>55</sup> siə<sup>33</sup> hu<sup>24</sup>. iŋ<sup>33</sup> kue<sup>21-53</sup> k<sup>h</sup>ɔi<sup>33</sup>  
 3SG PROG ask cow give cow help 3PL cross river  
 ‘It [the cat] is asking the cow to help it. They cross the river.’ [G3F1-3]

- (21) a. 伊人 就 艇 船 逗逗<sup>36</sup>  
i<sup>33</sup>naŋ<sup>11</sup> tɕiu<sup>11</sup> p<sup>h</sup>e<sup>24-11</sup> tɕuŋ<sup>55</sup> t<sup>h</sup>ik<sup>5</sup>t<sup>h</sup>ɔ<sup>55</sup>  
 3PL then paddle boat fun  
 ‘Then they paddle the boat for fun.’ [G1F4-16]
- b. 伊人 跳舞 清心 绝  
i<sup>33</sup>naŋ<sup>11</sup> t<sup>h</sup>iau<sup>21-53</sup>bu<sup>52</sup> tɕ<sup>h</sup>eŋ<sup>33</sup>sim<sup>33</sup> tɕɔ<sup>5</sup>  
 3PL dance content very  
 ‘They dance very contentedly.’ [G1F4-16]

CT Third Person Contextual ObjR Constructions can also be used to refer to inanimate objects. The examples below show 伊 *i<sup>33</sup>* ‘3SG’ referring to an advertisement in (22a) and a book in (22b).

- (22) a. 只个 船物 □ 巴士 顶畔 ... 伊个 广告  
 tɕia<sup>53</sup> tɕuŋ<sup>55</sup>-mue<sup>5</sup> na<sup>11</sup> pa<sup>33</sup>su<sup>24</sup> teŋ<sup>52-11</sup>pa<sup>55</sup> ... i<sup>33</sup> kai<sup>55</sup> kuaŋ<sup>11</sup>kau<sup>21</sup>  
 this.LW boat-thing LOC bus atop ... 3SG COP advertisement  
 ‘This boat thing is on a bus...It’s an advertisement.’ [G3F1-1]
- b. 只个 书 ... 伊 无 开 啊  
 tɕei<sup>55</sup> tɕu<sup>33</sup> ... i<sup>33</sup> bɔ<sup>11</sup> k<sup>h</sup>ui<sup>33</sup> a<sup>11</sup>  
 PDCO book ... 3SG NEG open PRT  
 ‘This [is] a book. It’s not open.’ [G1F1-12]

While both 伊人 *iŋ<sup>33</sup>* ‘3PL’ and 伊人 *i<sup>33</sup>naŋ<sup>11</sup>* ‘3PL’ are CT Third Person Plural Contextual ObjR Constructions (as seen in examples (20) and (21) respectively), at least for some speakers, there is some evidence that 伊 *i<sup>33</sup>* ‘3SG’ can also be used in plural contexts, as shown in (23).

---

<sup>36</sup> Note that I am using characters for this word from the [Taiwan Ministry of Education](#). This is in contrast with those used by Goh (2020, p. 343), 踢跔, and TCKnow LLC. (2015), 逗逗. The meaning of this common word is varied and includes things like sightseeing, playing, having fun, strolling, and leisure.

- (23) a. 二 人 伊 坐 在 象 顶  
 nɔ<sup>24-11</sup> naŋ<sup>55</sup> i<sup>33</sup> tɕɔ<sup>24</sup> tɔ<sup>11</sup> tɕ<sup>hi</sup>ɔ̃<sup>24</sup> teŋ<sup>52</sup>  
 two people 3PL sit LOC elephant atop  
 ‘Two people, they are sitting on top of an elephant.’ [G1M1-7]
- b. 人人 是 惊 ... 伊 无 惊 只个 □.  
 naŋ<sup>55-11</sup> naŋ<sup>55</sup> si<sup>11</sup> kiã<sup>33</sup> ... i<sup>33</sup> bɔ<sup>55-11</sup> kiã<sup>33</sup> tɕi<sup>52</sup> kai<sup>55</sup> nia<sup>33</sup>.  
 everyone COP fear ... 3PL NEG fear PDCO PRT.  
 伊 惊 老虎  
 i<sup>33</sup> kiã<sup>33</sup> lau<sup>24-11</sup> hɔu<sup>52</sup>  
 3PL fear tiger  
 ‘Everyone is afraid. They don’t fear this one. They fear the tiger.’  
 [G2F1US-18]
- c. 只个 苹果 伊 平样  
 tɕi<sup>52</sup> kai<sup>55</sup> p<sup>hej</sup> kuẽ<sup>52</sup> i<sup>33</sup> pẽ<sup>55-11</sup> iɔ̃<sup>11</sup>  
 PDCO apple 3PL same  
 ‘These are apples, they are the same.’ [G1F3-PS]

Some historical sources cite the third person *i*<sup>33</sup> being used in both singular and plural contexts including Fielde (1878, p. 12) and Dean (1841, p. 12). At the same time, Lim (1886, p. 42) notes *iŋ*<sup>33</sup> and *i*<sup>33</sup>*naŋ*<sup>55</sup> for Swatow Third Person Plural Contextual ObjR Constructions. As for current Third Person Plural Contextual ObjR Constructions in other varieties, Xu (2007, p. 74) wrote of *i*<sup>33</sup>*naŋ*<sup>55-11</sup> in Jieyang Chaozhou, Yeo (2011, p. 35) and Low (2014, p. 18) had *i*<sup>33</sup>*naŋ*<sup>11</sup> for Singapore Teochew, X. Chen (2003, p. 379) noted *i*<sup>33</sup>*laŋ*<sup>33</sup> for Johor Bahru Teochew. According to Shi (1999, pp. 302-3), there is variation amongst the dialects in Chaoshan as to what they use for Third Person Plural. Shi stated that the Plural Contextual ObjRs are historically formed by combining the singular forms with 人 *naŋ*<sup>55</sup> ‘person’. For the First and Second Person Contextual ObjR Constructions, this has resulted in single syllable words, as we see in Cambodian Teochew. For the Third Person, this has resulted in a two-syllable

construction, at least for some Chaoshan varieties (Shi, 1999) and the three varieties mentioned above.<sup>37</sup>

That at least some CT speakers use 伊人 *ij*<sup>33</sup> ‘3PL’ and 伊人 *i*<sup>33</sup>*naŋ*<sup>11</sup> ‘3PL’ likely reflects the different Chaoshan regions different families emigrated from. However, overall, there were not many tokens of CT Third Person Plural Contextual ObjR Constructions and so more data is needed to draw any more certain conclusions.

### 5.1.3.2 CT Deictic Contextual ObjR Constructions

Deictic contextual objR constructions<sup>38</sup> express a spatial location as it relates to the interlocutors, also known as spatial deixis (Croft, forthcoming, p. 59). Deictic contextual objR constructions stand alone, in contrast with deictic contextual attribute constructions (Croft, forthcoming, p. 63) which modify a referring expression. Cambodian Teochew has two Deictic Contextual ObjR Constructions. The CT Proximal Deictic Contextual ObjR Construction, 只个 *tɕi*<sup>52</sup>*kaɪ*<sup>55</sup> ‘PDCO’<sup>39</sup> makes reference to something that is close(r) to the speaker, as shown in (24). In (24a) and (24b), the Contextual ObjR 只个 *tɕi*<sup>52</sup>*kaɪ*<sup>55</sup> ‘PDCO’ refers to an item on the page of the story they are telling and thus it is retrievable from the non-linguistic discourse context. In (24c), the 只个 *tɕi*<sup>52</sup>*kaɪ*<sup>55</sup> ‘PDCO’ is referring to the referent ‘shell’, which was mentioned in a prior utterance.

---

<sup>37</sup> This is reflected in the characters as well. The First and Second Person Plural Contextual ObjR Constructions have one character (俺 and 汝, respectively) while the Third Person, even if it is only pronounced as one syllable, retains the two characters (伊人).

<sup>38</sup> These are Croft’s (forthcoming, p. 59) “demonstrative pronouns”.

<sup>39</sup> In glossing I have chosen to abbreviate the rather unruly CT Proximal Deictic Contextual ObjR Construction and CT Distal Deictic Contextual ObjR Construction terms to PDCO and DDCO, respectively.

- (24) a. 只个            乜个  
tei<sup>52</sup>kai<sup>55</sup>        mi<sup>?</sup>5kai<sup>55</sup>  
PDCO                what  
‘What is this?’ [G2M1-5]
- b. 我    唔    晓    但    只个  
ua<sup>52</sup>    m<sup>11</sup>    hiau<sup>24</sup>    tã<sup>21-53</sup>    tei<sup>52</sup>kai<sup>55</sup>  
1SG    NEG    know    say    PDCO  
‘I don’t know how to say this.’ [G2M1-5]
- c. “... 怎呢            伊    无    有    只个?”  
“...    tɛŋ<sup>52</sup>ni<sup>55-11</sup>        i<sup>33</sup>    bɔ<sup>11</sup>    u<sup>11</sup>    tei<sup>52</sup>kai<sup>55</sup>?”  
...    why                    3SG    NEG    have    PDCO  
“...why doesn’t he have this?” [G2F1US-17]

The CT Distal Deictic Contextual ObjR Construction, 许个 *hu<sup>52</sup>kar<sup>55</sup>* ‘DDCO’, makes reference to something further away from the speaker, as in (25).

- (25) a. 许个            有    生    加    □  
hu<sup>52</sup>kar<sup>55</sup>        u<sup>11</sup>    sɛ̃<sup>33</sup>    ke<sup>33</sup>    nɔ<sup>11</sup>  
DDCO                have    fruit    many    PRT  
‘That (one) has a lot of fruit.’ [G1F5US-23]
- b. 许个  
hu<sup>52</sup>kar<sup>55</sup>  
DDCO  
‘that (one)’ [G1F1-E]

The texts gathered for this dissertation had many examples of CT Presentational Constructions (see §7.2.2) with the use of the CT Proximal Deictic Contextual ObjR Construction, such as the examples in (26). In each utterance, the speaker uses 只个 *tei<sup>52</sup>kar<sup>55</sup>* ‘PDCO’ to refer to the picture itself. This usage is potentially a convention in Cambodian Teochew, or it could be a relic of the data gathering procedure/picture description task. Either way, it was frequently by five speakers, G1F1, G1F2, G1F3, G1M1, and G2M1. The strategy

was not found to be used by any of the G3 speakers. We could call the usage a topic comment strategy in which the topic is the picture and the comment is the rest of the utterance describing the picture.

- (26) a. 只个      伊      口              伊              脚头口  
tɕi<sup>52</sup>kai<sup>55</sup>      i<sup>33</sup>      nim<sup>24-11</sup>      i<sup>33</sup>              k<sup>h</sup>a<sup>33</sup>t<sup>h</sup>au<sup>11</sup>u<sup>33</sup>  
PDCO              3SG      hold              3SG.POSS      knee  
‘[In] this (one [the picture]), he holds his knees.’ [G1F1-12]
- b. 只个              口..      猫      死  
tɕi<sup>52</sup>kai<sup>55</sup>              um<sup>33</sup> ..      ŋiau<sup>33</sup>      uk<sup>5</sup>  
PDCO              INTJ..      cat      sleep  
‘[In] this (one [the picture]), a cat sleeps.’ [G1F2-14]
- c. 只个              旅行              船      在      海      中央  
tɕi<sup>52</sup>kai<sup>55</sup>              li<sup>52-24</sup>kiã<sup>55-11</sup>      tɕuŋ<sup>55</sup>      tɔ<sup>11</sup>      hai<sup>52-24</sup>      ta<sup>33</sup>uŋ<sup>33</sup>  
PDCO              travel              boat      LOC      sea      middle  
‘This (one [the picture]) [is] a cruise ship in the middle of the ocean.’  
[G1F3-15]
- d. 只个              厝      ...      只个              树  
tɕi<sup>52</sup>kai<sup>55</sup>              tɕ<sup>h</sup>u<sup>21</sup>      ...      tɕi<sup>52</sup>kai<sup>55</sup>              tɕ<sup>h</sup>iu<sup>11</sup>  
PDCO              house      ...      PDCO              tree  
‘This (one [the picture]) [is] a house... this (one [the picture]) [is] a tree.’  
[G1M1-7]
- e. 只个..              有      猪      有      马      有      厝  
tɕi<sup>52</sup>kai<sup>55</sup>..              u<sup>11</sup>      tu<sup>55</sup>      u<sup>11</sup>      be<sup>52</sup>      u<sup>11</sup>      tɕ<sup>h</sup>u<sup>21</sup>  
PDCO..              have      pig      have      horse      have      house  
‘[In] this (one [the picture]), there is a pig, there is a horse, there is a house.’ [G2M1-5]

In some ways, the instances of 只个 *tɕi<sup>52</sup>kai<sup>55</sup>* ‘PDCO’ in (26) resemble CT Deictic Contextual Attribute constructions, shown in (27). However, that the combination of 只 *tɕi<sup>52</sup>* ‘this’ and 个 *kai<sup>55</sup>* yields a Deictic Contextual ObjR Construction is clear because of the lack of tone sandhi on the *kai<sup>55</sup>* in (26). Conversely, in CT Deictic Contextual Attribute Constructions,

because the linker<sup>40</sup> word *kaɪ<sup>55</sup>* is in a modification relationship with the word that follows in the examples in (27), it undergoes tone sandhi.

- (27) a. 只 个 猫 坐 ...  
           *tɕi<sup>52</sup> kai<sup>55-11</sup> ŋiau<sup>33</sup> tɕɔ<sup>24</sup>* ...  
           this LW cat sit ...  
           ‘This cat sits.’ [G1F2-14]
- b. 伊 个 妘人 ... 口 许 个 丈夫 个 手  
       *i<sup>33</sup> kai<sup>55-11</sup> bəu<sup>52</sup> nən<sup>55</sup> ... nɪm<sup>24-11</sup> hu<sup>52</sup> kai<sup>55-11</sup> ta<sup>33</sup> pəu<sup>33</sup> kai<sup>55-11</sup> tɕiu<sup>52</sup>*  
       3SG.POSS LW wife ... hold that LW man LW hand  
       ‘His wife...holds that man’s hand.’ [G3F2-4]

The morphemes 只 *tɕi<sup>52</sup>* ‘this’ and 许 *hu<sup>52</sup>* ‘that’ cannot function alone as CT Deictic Contextual ObjR Constructions. The morpheme *kaɪ<sup>55</sup>* is necessary. Traditionally, *kaɪ<sup>55</sup>* is considered a generic classifier, though here we call it a linker because it has no clear semantic component (see Croft, forthcoming, p. 123). In Cambodian Teochew, Deictic + Classifier can be used to refer to an item without an overt ObjR if said object can be determined from context (see §6.2.1). This usage is likely how 只个 *tɕi<sup>52</sup>kaɪ<sup>55</sup>* and 许个 *hu<sup>52</sup>kaɪ<sup>55</sup>* came to become CT Deictic Contextual ObjR Constructions. However, there is no evidence that 只 *tɕi<sup>52</sup>* ‘this’ and 许 *hu<sup>52</sup>* ‘that’ can combine with any other classifiers to form CT Deictic Contextual ObjR Constructions. It is only the combination of 只 *tɕi<sup>52</sup>* ‘this’ and 许 *hu<sup>52</sup>* ‘that’ and the linker *kaɪ<sup>55</sup>* that produces these constructions. However, syllable contraction (see Myers & Li, 2009) gives rise to variations such as 只个 *tɕei<sup>55</sup>* ‘PDCO’ in (28) and 只个 *tɕiai<sup>55</sup>* ‘PDCO’ in (29). Note that the use of either of these forms varies based on speaker.

---

<sup>40</sup> See Chapter 6 for more information on the role of linkers in CT Modification Constructions.

- (28) a.  只个  鸡  
           tɕei<sup>55</sup>  kɔi<sup>33</sup>  
           PDCO  chicken  
           ‘This [is] a chicken.’ [G1F1-8]
- b.  只个  草          啊  
           tɕei<sup>55</sup>  tɕ<sup>h</sup>au<sup>52</sup>      a<sup>11</sup>  
           PDCO  grass          PRT  
           ‘This [is] grass.’ [G1F1-8]
- (29)  只个          鸟  
        tɕiai<sup>55</sup>      tɕiau<sup>52</sup>  
        PDCO          bird  
        ‘This [is] a bird.’ [G1F2-14]

The morphemes 只 *tɕi*<sup>52</sup> ‘this’ and 许 *hu*<sup>52</sup> ‘that’ can combine with other morphemes to create other types of CT Contextual ObjRs, as outlined in Table 5.3, organized by ontological type.

Proximal Deictic	Distal Deictic	Interrogative	Ontological type
只个 tɕi <sup>52</sup> kai <sup>55</sup> ‘PDCO’	许个 hu <sup>52</sup> kai <sup>55</sup> ‘DDCO’	底噢 (ti <sup>11</sup> )tiaŋ <sup>55</sup> ‘who’	person
		乜个 mi <sup>?</sup> kai <sup>55</sup> ‘what’	thing
只块 tɕi <sup>52</sup> kɔ <sup>21</sup> ‘here’	许块 hu <sup>52</sup> kɔ <sup>21</sup> ‘there’	底块 (ti <sup>11</sup> )kɔ <sup>21</sup> ‘where’	place
只阵 tɕi <sup>52</sup> tɕuŋ <sup>55</sup> ‘now’	许阵 hu <sup>52</sup> tɕuŋ <sup>55</sup> ‘then’	噢时 tiaŋ <sup>33</sup> si <sup>55</sup> ‘when’	time

Table 5.3: Types of CT Contextual ObjRs (adapted from Croft, forthcoming, p. 63)

## 5.2 Information Status Continuum

The information status continuum is a way of showing how languages communicate the information status of a referent (Croft, forthcoming, pp. 64-66). These distinctions in language tell us whether a referent is identifiable to the hearer, and if it is specific or nonspecific, real or imaginary, or generic. I present an adapted information status continuum and associated constructions in Table 5.4. The continuum ranges from easily identifiable referents at the top, to less and less readily identifiable referents, to solely reference to the type at the bottom.

Information Packaging			Constructions		
Information Properties				Article	Contextual ObjR
Identifiability	Referent status	Activation			
IDENTITY KNOWN §5.2.1	ACCESSIBLE IN DISCOURSE	active			anaphoric, zero
		semi-active		anaphoric (identity known)	
	ACCESSIBLE IN SHARED KNOWLEDGE	inactive		non-	
inferable		anaphoric (identity known)			
		Referent Categories			
IDENTITY UNKNOWN §5.2.2	REAL REFERENT	pragmatically specific		pragmatically specific	
		semantically specific	interrogative	pragmatically nonspecific; interrogative	
TYPE IDENTIFIABLE §5.2.3	NON-REAL “REFERENT”	nonspecific, various kinds:		(semantically) nonspecific	
		irrealis (context)			
		question (context)	conditional (context)		
		indirect negation (context)	comparative (context)		
		direct negation (context)	free choice (context)		
TYPE (not token/individual) §5.2.4		generic		generic	

Table 5.4: Information statuses and associated constructions (adapted from Croft, forthcoming, p. 66)

In Croft's version of Table 5.4, he includes terminology like "definite" and "indefinite" in regards to articles because English makes this distinction with the articles 'the' and 'a'. English Definite Articles are found under identity known constructions in Table 5.4, while English Indefinite Articles are used with the constructions in the identity unknown and type identifiable rows. There is not a definite/indefinite distinction in the articles in Cambodian Teochew, so we will not be using these terms here. Instead we will try to stick to terms related to the information packaging of referents.

The term ARTICLE is used to describe contextual forms combining with common objRs that express only information status (Croft, forthcoming, p. 65). Articles perform the situating subfunction of the modification function (see Chapter 6), but they are covered in this chapter because they relate to a referent's information status.

Croft stated that deictic contextual constructions (both objR and attribute) do not fit on the continuum (forthcoming, p. 67), though they are included in the Accessibility Scale in Table 5.5 (§5.2.1). According to Dryer (2014, p. e234), our deictic contextual constructions (his "demonstratives") by nature have a known identity because of their more specific meaning. Additionally, deictic contextual constructions possibly act as a historical source of some of the forms on the information status continuum (Croft, forthcoming, p. 67).

This section will generally look at ways of expressing information status in Cambodian Teochew. We will move down the continuum based on identifiability. §5.2.1 covers CT Identity Known Constructions at the top of Table 5.4, §5.2.2 highlights CT Identity Unknown Constructions, §5.2.3 focuses on CT Type Identifiable Constructions, and §5.2.4 showcases CT Generic ObjR Constructions.

### 5.2.1 CT Identity Known Constructions

The highest stage of the information status continuum is instances where the identity of the referent is known to both the speaker and hearer. For Cambodian Teochew, we can call these generally CT Identity Known Constructions.<sup>41</sup> These constructions can be divided further according to their accessibility, or activation, which describes how readily accessible a referent is to the hearer in the speaker's mind. There are three levels of accessibility of referents. Active (§5.2.1.1) is the most accessible, followed by semi-active (§5.2.1.2), with inactive being the least accessible. Referents can also be inferable based on the discourse context. Inactive and inferable referents are briefly covered in §5.2.1.3.

Croft (forthcoming, p. 71) presented an Accessibility Scale<sup>42</sup> of subtypes of referring phrases which I have adapted for Cambodian Teochew in Table 5.5. The higher an item is in the table, the greater the accessibility.

---

<sup>41</sup> Croft (forthcoming, p. 68) calls these “definite constructions”.

<sup>42</sup> Croft himself cites Givón (1983) and Ariel (1990).

Referring Phrase	Morphosyntactic expression
Zero Anaphora: $\emptyset$	zero
Anaphoric (Third Person) Contextual ObjR: <i>i<sup>33</sup></i> ‘3SG’, <i>ij<sup>33</sup>/i<sup>33</sup> naŋ<sup>11</sup></i> ‘3PL’	contextual objR
First/Second Person Contextual ObjR: <i>ua<sup>52</sup></i> ‘1SG’, <i>lu<sup>53</sup></i> ‘2SG’, etc.	contextual objR
Deictic Contextual ObjR: <i>tɕi<sup>52</sup> kaɪ<sup>55</sup></i> ‘PDCO’, <i>hu<sup>52</sup> kaɪ<sup>55</sup></i> ‘DDCO’	contextual objR
Deictic Contextual Attribute + ObjR: <i>tɕi<sup>52</sup> kaɪ<sup>55-11</sup> ŋiau<sup>33</sup></i> ‘this cat’	deictic contextual attribute + objR
Identity Known ObjR Phrase: <i>i<sup>33</sup> nɔu<sup>33</sup> kiã<sup>52</sup></i> ‘the child’, <i>kaɪ<sup>55-11</sup> lau<sup>24-11</sup> ho<sup>52</sup></i> ‘the tiger’	article + objR
proper name	proper objR

Table 5.5: Accessibility Scale for referring expressions in Cambodian Teochew (adapted from Croft, forthcoming, p. 71)

Per Croft (forthcoming), we find the following universal in regards to the accessibility of referents:

- (30) Shorter expressions (those higher on the scale) are used for higher accessibility referents than longer expressions. (Croft, forthcoming, p. 71)

Thus, the constructions at the top end of the Accessibility Scale are contextual objRs while those at the bottom end are common objRs that combine with an article or deictic contextual attribute (Croft, forthcoming, p. 72).<sup>43</sup>

---

<sup>43</sup> “Determiner” is a general term that encompasses articles and deictic contextual attribute constructions (Croft, forthcoming, p. 65) but we will avoid this term here.

### 5.2.1.1 Active (in focus)

Active referents are in focus, where a “discourse file has been opened and is at the center of hearer’s consciousness” (Croft, forthcoming, p. 68). From the Accessibility Scale for Cambodian Teochew in Table 5.5, the first two referring phrase types are used for active referents: Zero Anaphora and Anaphoric (Third Person) Contextual ObjR.

These two constructions are illustrated in the series of sentences in (31) and (32) below, from speakers G2M1 and G1F1, respectively. They each introduce the referent, 孛团 *nou<sup>33</sup>kiã<sup>52</sup>* ‘child’ and it becomes the most accessible referent in the discourse context. Afterwards, the two instances of 伊 *i<sup>33</sup>* ‘3SG’ in each example are CT Anaphoric Contextual ObjR Constructions, while the Øs represent CT Zero Anaphora Constructions. Each instance of *i<sup>33</sup>* ‘3SG’ and the Ø are referring to the previously mentioned, active referent, which is *nou<sup>33</sup>kiã<sup>52</sup>* ‘child’ in both examples.

- (31) 孛团 死 ... 伊 醒. 伊 肚困. Ø 欲 食 奶  
*nou<sup>33</sup>kiã<sup>52</sup> uk<sup>5</sup> ... i<sup>33</sup> tɕ<sup>h</sup>ẽ<sup>52</sup>. i<sup>33</sup> tɔu<sup>52</sup>k<sup>h</sup>uŋ<sup>21</sup>. Ø aĩ<sup>21</sup> tɕiaĩ<sup>5-2</sup> ni<sup>24</sup>*  
 child sleep ... 3SG wake 3SG hungry 3SG want eat milk  
 ‘A child sleeps... He wakes up. He’s hungry. [He] wants to eat milk.’ [G2M1-6]

- (32) 只个 孛团 在 死 ... Ø 死 在 铺 顶.  
*tɕei<sup>55</sup> nou<sup>33</sup>kiã<sup>52</sup> tɔ<sup>11</sup> i<sup>24</sup> ... Ø i<sup>24</sup> tɔ<sup>11</sup> p<sup>h</sup>ɔu<sup>33</sup> teŋ<sup>52</sup>.*  
 PDCO child PROG sleep ... 3sg sleep LOC bed atop  
 好 伊 妈妈.. 但 故事. 伊 听 伊 在 死 时候 啊  
*hɔ<sup>52</sup> i<sup>33</sup> ma<sup>33</sup>ma<sup>33</sup>.. tã<sup>21-53</sup> ku<sup>53</sup>suu<sup>11</sup>. i<sup>33</sup> t<sup>h</sup>iã<sup>33</sup> i<sup>33</sup> tɔ<sup>11</sup> i<sup>24</sup> si<sup>11</sup>hau<sup>11</sup> a<sup>33</sup>*  
 good 3SG mama.. tell story 3SG listen 3SG PROG sleep time PRT  
 ‘This [is] a child sleeping...[He] sleeps on top of the bed. Then his mother tells a story. He listens while he is sleeping.’ [G1F1-10]

ANAPHORIC usually means that the referent has been explicitly mentioned (Croft, forthcoming, p. 69), but it can also be determined from the nonlinguistic discourse context. A

CT Anaphoric Contextual ObjR Construction just needs to be highly accessible. The example in (33) is the first sentence uttered by speaker G1F1 for text #11, so there is no preceding referent for 伊 *i<sup>33</sup>* ‘3SG’. However, the speaker is pointing to the boy on the cover of the book while producing the utterance, so the referent becomes clear from the nonlinguistic discourse context.

- (33) 只个 伊 起 从 树  
       *tɕi<sup>52</sup>kai<sup>55</sup>*    *i<sup>33</sup>*    *kʰi<sup>52-24</sup>*    *tɕaŋ<sup>55-11</sup>*    *tɕ<sup>h</sup>iu<sup>11</sup>*  
       PDCO        3SG    go.up        CLF.TREE    tree  
       ‘[In] this (one [the picture]), he goes up a tree.’ [G1F1-11]

#### 5.2.1.2 *Semi-active (accessible)*

Semi-active referents are accessible in that they have an activated discourse file. They have been mentioned previously but are not currently the center of the hearer’s consciousness (Croft, forthcoming, p. 69). In Cambodian Teochew, semi-active referents can be indicated with a CT Anaphoric Article (Identity Known) Construction.

A referent might not be the center of a hearer’s consciousness (i.e. it’s less accessible) due to referential distance (how recently the referent was mentioned) and potential interference (other competing accessible referents) (Givón, 1983, pp. 13-15). Givon provides these two factors as a measure of a referent’s “topic continuity”.

This topic continuity can be demonstrated in the following series of examples from text #12 from speaker G1F1. In (34a), the referent 孳囡 *nɔu<sup>33</sup>kiã<sup>52</sup>* ‘child’ is introduced. This remains the active referent for the next utterance in (34b) and thus a CT Anaphoric Contextual ObjR Construction, 伊 *i<sup>33</sup>* ‘3SG’, is used.

- (34) a. 孛团      行      去      读书  
nɔu<sup>33</sup>kiã<sup>52</sup>      kiã<sup>55-11</sup> k<sup>h</sup>u<sup>21-53</sup> t<sup>h</sup>ak<sup>5-2</sup>tɕu<sup>33</sup>  
child      walk      go      read-book  
‘A child is walking to school.’ [G1F1-12]
- b. 伊      行      去      读书      啊  
i<sup>33</sup>      kiã<sup>55-11</sup> k<sup>h</sup>u<sup>21-53</sup>      t<sup>h</sup>ak<sup>5-2</sup>tɕu<sup>33</sup>      a<sup>33</sup>  
3SG      walk      go      read-book      PRT  
‘She walks to school.’ [G1F1-12]

Next, the speaker produces several more utterances, most of which are about the ‘child’, and thus speaker G1F1 uses 伊 *i<sup>33</sup>* ‘3SG’ as well as the CT Zero Anaphora Construction, Ø, to refer to her. Still, there are some competing referents, so the speaker feels the need to remind us who the active referent is by using the CT Anaphoric Article (Identity Known) Construction, 伊 *i<sup>33</sup>* which we will code as ‘IK’ for identity known, to modify the referent 孛团 *nɔu<sup>33</sup>kiã<sup>52</sup>* ‘child’ in (35). We can see that the CT Anaphoric (Third Person) Contextual ObjR 伊 *i<sup>33</sup>* ‘3SG’ shares the same form with the CT Anaphoric Article (Identity Known) Construction 伊 *i<sup>33</sup>* ‘IK’. According to Lü (1985), 伊 was a “demonstrative” form in the Chinese of the pre-Qin period (p. 17). Thus, in Cambodian Teochew (and other varieties) we are likely seeing this “demonstrative”/deictic form that has extended up the accessibility scale in Table 5.5 to grammaticalize into the CT Anaphoric (Third Person) Contextual ObjR, and at the same time has extended down the accessibility scale to become the CT Anaphoric Article (Identity Known) Construction. This is a typologically common phenomenon (Croft, forthcoming, p. 73-77).

- (35) 只个            伊      孛团            行行行  
 tɕi<sup>52</sup>kai<sup>55</sup>      i<sup>33</sup>      nɔu<sup>33</sup>kiã<sup>52</sup>      kiã<sup>55</sup> kiã<sup>55</sup> kiã<sup>55</sup>  
 PDCO            IK      child            walk, walk, walk<sup>44</sup>  
 ‘[In] this (one [the picture]), the child is walking a bit.’ [G1F1-12]

In the same story, a few utterances later, the ‘child’, who is still the active referent and can be referred to with 伊 *i<sup>33</sup>* ‘3SG’, sees another child, 弟弟 *didi* ‘boy’, in (36).

- (36) 只个            扶\*   簿   来   伊   睇   着   弟弟   在   读  
 tɕi<sup>52</sup>kai<sup>55</sup>      iau<sup>11</sup>   p<sup>h</sup>ou<sup>24</sup>   lai<sup>55</sup>   i<sup>33</sup>   tho<sup>52-24</sup>   tiɔ<sup>5</sup>   didi   to<sup>11</sup>   tak<sup>2</sup>  
 PDCO            open   book   come   3SG   see   RC   boy   PROG   study  
 ‘[In] this (one [the picture]), she opens the book and sees a boy studying.’  
 [G1F1-12]

Given that the story now has two competing referents, a girl and a boy, in future utterances the speaker needs to specify which one she is talking about, which in the case of (37) is the 弟弟 *didi* ‘boy’.<sup>45</sup>

- (37) 弟弟   伊   在   睇..            睇   着..   簿  
didi   i<sup>33</sup>   to<sup>11</sup>   t<sup>h</sup>ɔi<sup>52-24</sup> ..      t<sup>h</sup>ɔi<sup>52-24</sup>   tiɔ<sup>2</sup> ..   p<sup>h</sup>ou<sup>24</sup>  
boy   3SG   PROG   look..            look   RC..   book  
 ‘The boy, he is looking at a book.’ [G1F1-12]

The referent in the start of the very next utterance in (38) is the first ‘child’. While she was previously able to be referred to as 孛团 *nɔu<sup>33</sup>kiã<sup>52</sup>* ‘child’, like in (34a), now that there are two ‘children’ in the story, the speaker needs to be more specific, so she says 姐姐 *tɕe<sup>52-24</sup>tɕe<sup>52</sup>*

<sup>44</sup> Tone sandhi is expected here on the two initial *kiã<sup>55</sup>* syllables yet all three surface in citation tone.

<sup>45</sup> Note that even though the identity is still known for the ‘boy’, CT Anaphoric Article (Identity Known) Constructions are not required. And the lack of such a marker should not be taken to mean that the referent’s identity is unknown (Dryer, 2014, p. e234).

‘girl’. Even though the speaker hasn’t used the term 姐姐 *tɕe<sup>52-24</sup>tɕe<sup>52</sup>* ‘girl’ prior to the utterance in (38), the ‘girl’ is still a semi-active referent in the discourse context, so the speaker is able to use the CT Anaphoric Article (Identity Known) Construction, 伊 *i<sup>33</sup>* ‘IK’.<sup>46</sup>

- (38) 伊      姐姐              是      □      伊      簿      来      了  
*i<sup>33</sup>*      *tɕe<sup>52-24</sup>tɕe<sup>52</sup>*      si<sup>11</sup>      nim<sup>24-11</sup> i<sup>33</sup>      p<sup>h</sup>ou<sup>24</sup> lai<sup>55</sup>      ou<sup>11</sup>  
IK      girl              be      hold      IK      book      come      already  
‘The girl is holding her book.’ [G1F1-12]

伊 *i<sup>33</sup>* ‘IK’ is not the only CT Identity Known Construction used for semi-active referents. Another construction, known as the “bare classifier” construction in the literature (Cheng & Sybesma, 2005; Li & Bisang, 2012; Simpson, Soh, & Nomoto, 2011; Wang Jian, 2015), can also indicate identity known referents in some languages. As will be shown in Chapter 6, classifiers are traditionally used in CT Enumeration Modification Constructions with the canonic word order of Enumeration/Deictic Contextual Attribute + Classifier + ObjR. The “bare classifier” constructions contain only the classifier and the objR. These constructions are distinct from “bare” objRs which contain only the head objR.<sup>47</sup> We show how this distinction is especially relevant when discussing CT Generic ObjR Constructions in §5.2.4.

---

<sup>46</sup> Similar anaphoric usages of 伊 *i<sup>33</sup>* have been noted for Jieyang (Xu, 2007, pp. 76-77) and Singapore (Low, 2014, pp. 23-24) varieties.

<sup>47</sup> Per Croft (forthcoming), the head of a phrase is “the most contentful word that most closely denotes the same function as the phrase (or clause) as a whole” (p. 32). While “noun” in Croft (forthcoming) represents the head of a referring phrase when the head is an object, Croft uses the phrase “referent expression” to refer to the head of a referring phrase that is not necessarily an object (p. 39). Thus in the clause ‘hiking is fun’, ‘hiking’ is action reference and is a referent expression.

When the “bare classifier” is used in topic position in Cambodian Teochew, it acts as a CT Identity Known Construction.<sup>48</sup> In the examples in (39), the identity of the relevant referents is known, as they were mentioned in a prior utterance.

- (39) a.  只个      隻            鸡      伊    企      在      篱笆      顶  
 tɕi<sup>52</sup>kai<sup>55</sup>    tɕiaŋ<sup>2-5</sup>            kɔi<sup>33</sup>    i<sup>33</sup>    k<sup>h</sup>ia<sup>24</sup>    tɔ<sup>11</sup>    li<sup>11</sup>pa<sup>55</sup>    teŋ<sup>52</sup>  
 PDCO        CLF.IK            rooster 3SG stand    LOC    fence    atop  
 ‘[In] this (one [the picture]), the rooster, he is standing on top of the fence.’ [G1F3-15]
- b.  伊      爬      起      来      就      隻      狗..  
 i<sup>33</sup>    peŋ<sup>2</sup>    k<sup>h</sup>i<sup>52</sup>    lai<sup>55</sup>    tɕiu<sup>11</sup>    tɕiaŋ<sup>2-5</sup>    kau<sup>52</sup>..  
 3SG climb up    come then CLF.IK dog..  
 被      在      伊            头      顶  
 p<sup>h</sup>ɔŋ<sup>2</sup>    tɔ<sup>11</sup>    i<sup>33</sup>            t<sup>h</sup>au<sup>55-11</sup>    teŋ<sup>52</sup>  
 hang    LOC    3SG.POSS    head    atop  
 ‘He gets up and then the dog hangs on top of his head.’ [G1F1-11]

Such “bare classifiers” only certainly indicate an identity known referent when in topic position. In other positions, the interpretation is contextual. While (40) gives an identity known interpretation because the 狗 *kau*<sup>52</sup> ‘dog’ has already been mentioned in the discourse context, the 狗 *kau*<sup>52</sup> ‘dog’ and 树 *tɕ<sup>h</sup>iu*<sup>11</sup> ‘tree’ in (41) are new referents with an unknown identity.

---

<sup>48</sup> An anonymous reviewer noted that the topic position in (39a) appears to be filled by 只个 *tɕi*<sup>52</sup>*kai*<sup>55</sup> ‘PDCO’. Indeed, in §5.1.3.2, we have called this convention in Cambodian Teochew the topic comment strategy. However, that is not to say that 隻鸡 *tɕiaŋ*<sup>2-5</sup>*kɔi*<sup>33</sup> ‘CLF.IK rooster’ is not also a topic. We follow LaPolla (1990, p. 143) to call 只个 *tɕi*<sup>52</sup>*kai*<sup>55</sup> ‘PDCO’ the primary topic (the picture being described) and 隻鸡 *tɕiaŋ*<sup>2-5</sup>*kɔi*<sup>33</sup> ‘CLF.IK rooster’ (the rooster in the picture) the secondary topic that itself is part of the assertion on the primary topic.

- (40) 二 隻 水鸡 合 隻 狗  
 nɔ<sup>24-11</sup> tɕiaŋ<sup>2-5</sup> tɕui<sup>52-24</sup> kɔi<sup>33</sup> kaŋ<sup>2</sup> tɕiaŋ<sup>2-5</sup> kau<sup>52</sup>  
 two CLF.ANIMAL frog with CLF.IK dog  
 ‘Two frogs with the dog.’ [G1F1-11]

- (41) a. 有-有 只 个 兄人 在 挈 隻 狗 行  
 u<sup>11</sup>-u<sup>11</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> hiã<sup>33</sup> naŋ<sup>55</sup> tɔ<sup>11</sup> k<sup>h</sup>iɔŋ<sup>5</sup> tɕiaŋ<sup>2-5</sup> kau<sup>52</sup> kiã<sup>55</sup>  
 have- have this LW young.man PROG take CLF dog walk  
 ‘There is this young man taking a dog for a walk.’ [G1F4-16]
- b. 只个 伊 起 丛 树  
 tɕi<sup>52</sup>kai<sup>55</sup> i<sup>33</sup> k<sup>h</sup>i<sup>52-24</sup> tɕaŋ<sup>55-11</sup> tɕ<sup>h</sup>iu<sup>11</sup>  
 PDCO 3SG go.up CLF tree  
 ‘[In] this (one [the picture]), he goes up a tree.’ [G1F1-11]

In the examples in (39), the classifier is a CT Identity Known Construction. When the morpheme 个 *kaŋ<sup>55</sup>* is used this way, it is an Article, thus resulting in another CT Anaphoric Article (Identity Known) Construction. This type of usage was first noted in McFarland (2017, p. 24) with the example in (42).<sup>49</sup> More examples are provided in (43) from the same speaker, G2F1US, who uses this construction quite prolifically, perhaps due to influence from English. Nonetheless, the construction is still used by the primary speaker consultants, as shown in (44). All of the referents in (42-44) that follow 个 *kaŋ<sup>55</sup>* ‘IK’ are semi-active referents. They have been previously mentioned yet are not the center of the hearer’s consciousness due to other intervening referents or referential distance.

- (42) “个 虫, 伊 无 平样 俺”  
 “kai<sup>55-11</sup> t<sup>h</sup>aŋ<sup>55</sup>, i<sup>33</sup> bɔ<sup>55-11</sup> pɛ<sup>55-11</sup>i<sup>5</sup> naŋ<sup>52</sup>”  
 IK worm 3SG NEG same 1PL”

<sup>49</sup> Note that the example has been updated from McFarland (2017) to correct errors and include tone markings.

“The worm, he is not the same as us.” [G2F1US-17]

- (43) a. 伊 但 “个 铁 门 真 大  
 i<sup>33</sup> tā<sup>21</sup> “kai<sup>55-11</sup> tiŋ<sup>2-5</sup> muŋ<sup>55</sup> tɕiŋ<sup>33</sup> tua<sup>11</sup>  
 3SG say “IK iron door very big  
 ‘He says “The iron door is very big.” [G2F1US-20]
- b. 个 鸟 在 中央.. 中央 垃圾  
kai<sup>55-11</sup> tɕiau<sup>52</sup> tɔ<sup>11</sup> ta<sup>33</sup>uŋ<sup>33</sup>.. ta<sup>33</sup>uŋ<sup>33</sup> naŋ<sup>5</sup>sap<sup>2</sup>  
IK bird LOC middle.. middle garbage  
 ‘The bird is in the middle of garbage.’ [G2F1US-PS]
- c. 个 壳  
kai<sup>55-11</sup> k<sup>h</sup>ak<sup>2</sup>  
IK shell  
 ‘The shell’ [pointing to the shell] [G2F1US-17]
- d. 个 老虎 在 死  
kai<sup>55-11</sup> lau<sup>24-11</sup>həu<sup>52</sup> tɔ<sup>11</sup> ɿ<sup>24</sup>  
IK tiger PROG sleep  
 ‘The tiger is sleeping.’ [G2F1US-18]
- (44) a. 个 大夫 拍 伊 妘  
kai<sup>55-11</sup> ta<sup>33</sup>pəu<sup>33</sup> p<sup>h</sup>aɿ<sup>2-5</sup> i<sup>33</sup> bəu<sup>52</sup>  
IK man hit 3SG.POSS wife  
 ‘The man hits his wife.’ [G3F3-4]
- b. 个 孳团 好 死  
kai<sup>55-11</sup> nəu<sup>11</sup>kiã<sup>52</sup> hɔ<sup>52-24</sup> uk<sup>5</sup>  
IK child good sleep  
 ‘The child sleeps well.’ [G1F4-16]

### 5.2.1.3 Inactive and Inferable

Inactive referents are ones that are identifiable to both the speaker and hearer but they haven’t yet been activated in the current discourse context (Croft, forthcoming, p. 69). A referent like ‘the president’ is an example of an inactive referent in English, assuming both the

speaker and hearer have a shared knowledge of who the current president is (and what country they are talking about), and if she hasn't been previously mentioned.

There is a similar yet slightly different information packaging for inferable referents, referents which have not been activated in the discourse context but their identity can be inferred based on descriptors or shared world knowledge (Croft, forthcoming, p. 70, citing Prince, 1992). Inactive and inferable referents are often grouped together, for example by Dryer (2014).

There were a few examples of these types of referents accessible in the shared knowledge in the current data set. We already showed in §5.1.2 how the shared discourse context can give special meanings to CT Proper ObjRs like 金塔 *kim<sup>33</sup>t<sup>h</sup>aŋ<sup>2</sup>* ‘Phnom Penh’ in (45), where that referent as well as 高棉 *kau<sup>33</sup>mian<sup>55</sup>* ‘Cambodia’ are examples of inactive referents. Similarly, in (46), shared knowledge gives rise to the interpretation of Mandarin from 国语 *kɔk<sup>2-5</sup>gu<sup>52</sup>* ‘country-language’, Teochew from 唐人 *tun<sup>55-11</sup>-naŋ<sup>55</sup>* ‘Chinese-people’, and Khmer from 番人 *huan<sup>33</sup>-naŋ<sup>55</sup>* ‘foreign-people’.

- (45) 我 来 ... 金塔 高棉  
 ua<sup>52</sup> lai<sup>55-11</sup> ... kim<sup>33</sup>t<sup>h</sup>aŋ<sup>2</sup> kau<sup>33</sup>mian<sup>55</sup>  
 1SG come ... Phnom Penh Cambodia (Khmer)  
 ‘I come to Phnom Penh, Cambodia.’ [G2M1-5]

- (46) 只个 Ø 知 但 国语 木瓜,  
 tɕai<sup>55</sup> Ø tɕai<sup>33</sup> tã<sup>21-53</sup> kɔk<sup>2-5</sup>gu<sup>52</sup> *mùguā*,  
 PDCO 1SG know COMP Mandarin papaya,  
 唐人 个 𠵼𠵼, □ 番人 𠵼𠵼  
tun<sup>55-11</sup>-naŋ<sup>55</sup> kai<sup>55-11</sup> lo<sup>11</sup>hɔŋ<sup>55</sup>, a<sup>33</sup> huan<sup>33</sup>-naŋ<sup>55</sup> *lhɔŋ*  
Teochew-people LW papaya INTJ Khmer-people papaya

‘This (one [the fruit]), I know that Mandarin [is] *mùguā*, [is] Teochew people’s *lɔ<sup>11</sup>hɔŋ<sup>55</sup>*, and Khmer *lhɔŋ*.’ [G1F5US-23]

Other examples of inactive referents include 日 *dʒik<sup>5</sup>* ‘sun’ (47), 月娘 *gueŋ<sup>5-2</sup>niõ<sup>55</sup>* ‘moon’ (48), and 神 *siŋ<sup>33</sup>* ‘god’ (49).

- (47) a. 圓 肖 日 平样  
*ĩ<sup>55</sup> siau<sup>24</sup> dʒik<sup>5</sup> pẽ<sup>55-11</sup>iõ<sup>11</sup>*  
 circle looks like sun same  
 ‘[The] circle looks like [the] sun.’ [G1F1-PS]
- b. 日 落 死  
*dʒik<sup>5</sup> loŋ<sup>5-2</sup> si<sup>52</sup>*  
 sun descend die  
 ‘[The] sun sets.’ [G3F1-3]
- c. 日 又 起 啊  
*dʒik<sup>5</sup> iu<sup>11</sup> kʰi<sup>52</sup> a<sup>11</sup>*  
 sun again ascend PRT  
 ‘[The] sun rises again.’ [G1F4-16]
- (48) a. 暝天 顶下 星 数. 月娘 啊是 起  
*mẽ<sup>55</sup>kua<sup>21</sup> tʰi<sup>33</sup>teŋ<sup>52</sup> e<sup>24-11</sup> tɕ<sup>h</sup>ẽ<sup>33</sup> tɕoi<sup>11</sup>. gueŋ<sup>5-2</sup>niõ<sup>55</sup> a<sup>11</sup>si<sup>11</sup> kʰi<sup>52</sup>*  
 night sky below star many. moon also ascend  
 ‘There are many stars below [the] night sky. [The] moon also rises.’  
 [G1F4-16]
- b. 只个 月娘. 只个 飞机  
*tɕei<sup>55</sup> bueŋ<sup>5-2</sup>niõ<sup>55</sup>. tɕei<sup>55</sup> pue<sup>33</sup>ki<sup>33</sup>*  
 PDCO moon PDCO airplane  
 ‘This [is the] moon. This [is an] airplane.’ [G1F1-10]
- (49) 每 年 谢 神 ...  
*mue<sup>52-24</sup> ni<sup>55</sup> sia<sup>11</sup> siŋ<sup>33</sup> ...*  
 every year thank god ...  
 ‘[I] thank god every year...’ [G1M2-24]

Based on the examples provided in this section, there do not appear to be any unique strategies for introducing inactive or inferable referents into the discourse context. In fact, (48b) shows the same strategy, a bare referent expression, for presenting an inactive referent 月娘 *bueŋ<sup>5-2</sup>niŋ<sup>55</sup>* ‘moon’ as for presenting an identity unknown referent (see the next section), 飞机 *pue<sup>33</sup>ki<sup>33</sup>* ‘airplane’.

### 5.2.2 CT Identity Unknown Constructions

So far we have covered the identity known portion at the upper end of the information status continuum from Table 5.4. Moving down we find constructions with referents whose identity is unknown to at least one of the speaker or hearer. In these constructions, there is a real referent that is semantically specific. We will call these CT Identity Unknown Constructions.

While all identity unknown constructions are semantically specific, Dryer (2014) made a division as to a referent’s pragmatic specificity. Though difficult to define, typically if a referent is mentioned and the speaker goes on to talk about the referent, then it is pragmatically specific and if the speaker does not go on to talk about said referent, then it is pragmatically unspecific (Dryer, 2014, p. e236). Dryer argued for this distinction because some languages have special constructions to introduce those referents that have “some sort of prominence in the discourse”, i.e. those that are pragmatically specific (Dryer, 2014, p. e237).

There are CT Identity Unknown Constructions which indicate a referent is specific with an identity unknown through the use of 蜀 *tɕek<sup>5-2</sup>* ‘one’. ‘One’ is often a source for specific identity unknown articles (Croft, forthcoming, p. 81; Heine & Kuteva, 2002, p. 220). In (50a), which is the first utterance in text #18, we see speaker G2F1US introduce the 老虎 *lau<sup>24-11</sup>hou<sup>52</sup>* ‘tiger’. Then a few utterances later in (50b) the 小狐狸 *xiǎo húlí* ‘fox’ is introduced. The same

speaker similarly introduces the 鱼 *hu*<sup>55</sup> ‘fish’ in (50c) from text #19. These three referents are pragmatically specific as they are the main characters in the respective stories. Another CT Identity Unknown Construction is shown in (50d), where the 柴 *tɕ<sup>h</sup>a*<sup>55</sup> ‘wood’ is introduced for the first time with 蜀 *tɕek*<sup>5-2</sup> ‘one/a’. The 柴 *tɕ<sup>h</sup>a*<sup>55</sup> ‘wood’ is also pragmatically specific as the speaker continues talking about it in the next utterance. The examples in (50) are all CT Presentational Constructions using the have possessive strategy (see §7.2.2). Speaker G1F2 similarly introduces the 猫 *ɲiau*<sup>33</sup> ‘cat’ in (51), who is also the main character of the story and thus is pragmatically specific.

- (50) a. ̀ɿŋ 底.. 有 蜀 隻.. 老虎  
 pɔ<sup>33</sup>loi<sup>52</sup> tɔi<sup>52</sup>.. u<sup>11</sup> tɕek<sup>5-2</sup> tɕiaŋ<sup>2-5</sup>.. lau<sup>24-11</sup>hou<sup>52</sup>  
 jungle inside..have one CLF.ANIMAL.. tiger  
 ‘Inside the jungle, there is a tiger.’ [G2F1US-18]
- b. 有 蜀 隻.. □ 小狐狸 ...  
 u<sup>11</sup> tɕek<sup>5-2</sup> tɕiaŋ<sup>2-5</sup>.. um<sup>33</sup> xiǎo húlí ...  
 have one CLF.ANIMAL.. INTJ fox ...  
 ‘There is a fox...’ [G2F1US-18]
- c. 老-古早 伊 伊 有 蜀,蜀 礼 鱼  
 lau<sup>24-11</sup>-kou<sup>52</sup>tɕa<sup>52</sup> i<sup>33</sup> i<sup>33</sup> u<sup>11</sup> tɕek<sup>5-2</sup>,tɕek<sup>5-2</sup> loi<sup>52-24</sup> hu<sup>55</sup>  
 old-past 3SG 3SG have one, one special fish  
 ‘In the olden days, there was a special fish.’ [G2F1US-10]
- d. 有 蜀 柴.. 蜀,蜀 块 啊  
 u<sup>11</sup> tɕek<sup>5-2</sup> tɕ<sup>h</sup>a<sup>55</sup>.. tɕek<sup>5-2</sup>,tɕek<sup>5-2</sup> kɔ<sup>21</sup> a<sup>55</sup>  
 have one wood.. one, one CLF-PIECE PRT  
 ‘There is one wood..a piece of wood.’ [G1F1-11]
- (51) 蜀 猫 死 啊  
tɕek<sup>5-2</sup> ɲiau<sup>33</sup> uk<sup>5</sup> a<sup>33</sup>  
one cat sleep PRT  
 ‘A cat sleeps.’ [G1F2-14]

While the five referents in (50) and (51) are all pragmatically specific, it is unclear if the use of 蜀 *tɕek*<sup>5-2</sup> ‘one/a’ necessitates this of a referent. It is certainly not the only way to introduce a pragmatically specific referent into the discourse. Referents are often introduced without any pre-modification as in (52), where each of the referents is pragmatically specific and mentioned in the next utterance of its respective text. The example in (52a) shows that the use of 蜀 *tɕek*<sup>5-2</sup> ‘one/a’ is not required for pragmatically specific referents in CT Presentational Constructions using the have possessive strategy.

- (52) a. 有 鸟 口 顶 树 顶畔  
 u<sup>24</sup> tɕiau<sup>52</sup> na<sup>11</sup> teŋ<sup>52-11</sup> tɕ<sup>h</sup>iu<sup>11</sup> teŋ<sup>52-11</sup> paĩ<sup>55</sup>  
 have bird LOC atop tree atop  
 ‘There are birds on top of a tree.’ [G3F1-3]
- b. ...Ø 睇见 獠\*鼠 口 土墙 顶畔  
 ...Ø t<sup>h</sup>ɔĩ<sup>52-11</sup> kiaŋ<sup>21-53</sup> ŋiau<sup>11</sup> tɕ<sup>h</sup>u<sup>52</sup> na<sup>11</sup> t<sup>h</sup>ɔu<sup>11</sup> tɕ<sup>h</sup>iɔ<sup>55</sup> teŋ<sup>52-11</sup> paĩ<sup>55</sup>  
 ...3SG see mouse LOC wall atop  
 ‘... [he] sees a mouse on top of a wall.’ [G3F1-3]
- b. ...伊 睇见 猫 口 ... 衫裤 个 篮  
 ...i<sup>33</sup> t<sup>h</sup>ɔĩ<sup>52-11</sup> kiaŋ<sup>21-53</sup> ŋiau<sup>33</sup> na<sup>11</sup> ... sã<sup>33</sup>-k<sup>h</sup>ɔu<sup>21</sup> kai<sup>55-11</sup> na<sup>55</sup>  
 ...3SG see cat LOC ... clothes LW basket  
 ‘... he sees a cat in a clothes basket.’ [G3F1-2]
- c. Ø 睇 猫  
 Ø thɔĩ<sup>52-24</sup> ŋiau<sup>33</sup>  
 3SG see cat  
 ‘[He] sees a cat.’ [G2M1-6]

In the examples in (50-52), the identity of the referent was unknown to the hearer. Another type of identity unknown construction is an interrogative, where the identity of the referent is unknown to the speaker. This is the case for the following examples. In (53), the character in the story is asking who woke him. He, as the speaker, does not know the identity

of the character who woke him. In (54), the character in the story is asking the question and again doesn't know the identity of the answer, the name of the other character.

(53) 伊 但 “底 来 唔 我 醒?”  
 i<sup>33</sup> tā<sup>21</sup> “ti<sup>11</sup>tian<sup>55</sup> lai<sup>55</sup> mui<sup>21-53</sup> ua<sup>52</sup> tɕ<sup>h</sup>ẽ<sup>52</sup>?”  
 3SG say who come NEG.scared 1SG wake”  
 ‘He says “Who is not afraid to come and wake me?”’ [G2F1US-18]

(54) “汝 名 乜个?”  
 “lu<sup>52</sup> miã<sup>55-11</sup> mi<sup>?</sup>kai<sup>55</sup>?”  
 2SG be.named what  
 “‘What are you named?’” [G2F1US-19]

The 底 *ti<sup>11</sup>tian<sup>55</sup>* ‘who’ and 乜个 *mi<sup>?</sup>kai<sup>55</sup>* ‘what’ in (53-54) are examples of CT Interrogative Contextual ObjR Constructions. Such interrogative contextual objR constructions are thought to be the diachronic source of another construction, which we’ll call indefinite<sup>50</sup> contextual objRs (Haspelmath, 1997, p. 27). Haspelmath (1997) presented nine categories of indefinite contextual objRs. Two of these are for real referents (specific) and will be covered here while the remaining seven are for non-real referents (non-specific) and will be covered in §5.2.3.

Haspelmath (1997) divided identity unknown constructions differently from Dryer (2014). With specific known referents, the speaker knows the identity of the real-world referent, but the hearer does not, while with specific unknown referents, neither the speaker nor hearer know the identity of the real-world referent (Croft, forthcoming, p. 84-85, citing Haspelmath 1997). To my knowledge there were not any examples of CT Specific Known

---

<sup>50</sup> The use of the term indefinite might be unavoidable because it is currently the only term that covers both the identity unknown and type identifiable sections of Table 5.4.

Contextual ObjRs in the current data set. Some CT Specific Unknown Contextual ObjRs are shown in (55). These are all specific and unknown to both the speaker and the hearer. For example, in (55a), a real something is being cooked, but the speaker, G1F1-10, does not know what it is.

- (55) a. 只个, 伊 在 酸 乜个  
 tɕei<sup>55</sup>, i<sup>33</sup> tɔ<sup>11</sup> suŋ<sup>33</sup> miʔ<sup>55</sup>kai<sup>55</sup>  
 PDCO, 3SG PROG cook something  
 ‘[In] this (one [the picture]), he is cooking something.’ [G1F1-10]
- b. 伊 睇 着 只个. □ 簿 啊是 乜个  
 i<sup>33</sup> thoŋ<sup>52-24</sup> tɔʔ<sup>2</sup> tɕi<sup>52</sup>kai<sup>55</sup>. u<sup>33</sup> p<sup>h</sup>ɔu<sup>24</sup> a<sup>11</sup>si<sup>24</sup> miʔ<sup>55</sup>kai<sup>55</sup>  
 3SG see RC PDCO. INTJ book or something  
 ‘He saw this (one). [It’s] a book or something.’ [G1F1-12]
- c. Ø 行 去 块 □  
 Ø kiã<sup>55-11</sup> k<sup>h</sup>u<sup>21-53</sup> k<sup>h</sup>ɔ<sup>21</sup> nia<sup>11</sup>  
 3SG walk go where PRT  
 ‘[She] walks somewhere.’ [G1F1-10]

While in some languages, like English, indefinite contextual objRs can take many forms like *someone*, *anyone*, *no one*, etc., this is not the case for Cambodian Teochew. The CT Indefinite Contextual ObjR Constructions shown here and in the next section are identical in form to CT Interrogative Contextual ObjR Constructions, which can be seen by comparing the Contextual ObjR in (54) to those in (55a-b).<sup>51</sup>

---

<sup>51</sup> Note that Khmer Interrogative Contextual ObjR Constructions and Khmer Indefinite Contextual ObjR Constructions also share the same form (Huffman, 1967), as shown below.

- (i.a) nih qwəy?  
 this what  
 ‘What is this?’ (Huffman, 1967, p. 160)
- (i.b) baə look cəŋ tiñ qwəy ...  
 if you wish buy something ...

### 5.2.3 CT Type Identifiable Constructions

The next section on the information status continuum from Table 5.4 is referents that are only type identifiable. These are non-real or non-specific referents. Thus they are non-identifiable and are associated with non-real situations including desires, hopes, and expectations (Croft, forthcoming, p. 78). In Cambodian Teochew, the use of Interrogative Contextual ObjRs as CT Indefinite Contextual ObjRs can indicate referents of this nature. Haspelmath (1997) presented seven types of these non-specific objRs (summarized by Croft, forthcoming, pp. 85-87).

With irrealis contextual objRs, the referents exist in contexts such as a wish, desire, or command (Croft, forthcoming, p. 85). Referents of this nature include those in future contexts, as well as contexts with non-indicative modality (Haspelmath, 1997, p. 40). There were no tokens of this type in my data set but the Jieyang Irrealis Contextual ObjR was found in Xu (2007), in (56) using the Generic ObjR 人 *naŋ*<sup>55</sup> ‘person’.<sup>52</sup> The ‘someone’ only exists in the hypothetical world in which the mirror actually gets damaged. More data is needed to see if these same types of constructions are found in Cambodian Teochew.

(56)	heŋ <sup>11</sup> tsai <sup>35</sup>	sia <sup>35-21</sup> hue <sup>35-21</sup>	ti <sup>11</sup> uã <sup>33</sup>	hoŋ <sup>2</sup> tsap <sup>5</sup> ,	buŋ <sup>2</sup> tsek <sup>5-2</sup> tiã <sup>11</sup>	to <sup>35-21</sup>
	now	society	order	complex	maybe	at
	mou <sup>53-35</sup> sẽ <sup>33</sup>	ts <sup>h</sup> ej <sup>55-11</sup> k <sup>h</sup> uaŋ <sup>213-53</sup>	e <sup>55</sup> ,	lu <sup>53</sup>	tiõ <sup>33</sup>	mo <sup>55-11</sup> t <sup>h</sup> oŋ <sup>2</sup> -ts <sup>h</sup> ia <sup>33</sup>
	some	circumstance	under	2SG	CLF	motor-bike
	puŋ <sup>2</sup> -meŋ <sup>55-11</sup> -puŋ <sup>2</sup> -peŋ <sup>5</sup>	k <sup>h</sup> eŋ <sup>2</sup> -	i <sup>33</sup>	p <sup>h</sup> ua <sup>213-53</sup>	ni <sup>55</sup> -	k <sup>h</sup> u <sup>213-53</sup> hueŋ <sup>25-2</sup> tsia <sup>11</sup>
	not-bright-not-white	PASS	3SG	puncture-	RVC	or

---

‘If you want to buy something...’ (Huffman, 1967, p. 237)

<sup>52</sup> According to Haspelmath (1997, p. 307), 人 *rén* ‘person; someone’ is also frequently used as an indefinite in Mandarin Chinese.

tā<sup>213-53</sup> k<sup>h</sup>eŋ<sup>2</sup> nan<sup>55-11</sup> tsuaŋ<sup>35-21-</sup> tiau<sup>11</sup> kai<sup>55-11</sup> kiā<sup>213</sup> ...  
 say PASS person hit- RVC CLF mirror

‘The social situation is very complex. You never know when your motor-bike will suddenly get a puncture or will have the mirror damaged by someone.’ (Xu, 2007, p. 207)

With question contextual objRs, the referent falls under the scope of the interrogative (Croft, forthcoming, p. 85). No examples were found in the current data sample. There is a Jieyang construction with the Generic ObjR 人 *nan*<sup>55</sup> ‘person’, as in (57).

- (57) a. u<sup>55-11</sup> nan<sup>55-11</sup> k<sup>h</sup>ui<sup>33</sup> ts<sup>h</sup>ia<sup>33</sup> san<sup>213</sup> neŋ<sup>53-213</sup> bo<sup>55-11</sup>  
 have people drive vehicle send 2PL not.have<sup>53</sup>  
 ‘Is anyone driving you?’ (Xu, 2007, p. 77)
- b. hio<sup>53-35</sup> lai<sup>35</sup> u<sup>35-21</sup> nan<sup>55</sup> bo<sup>55-11</sup>  
 that home have people not.have  
 ‘Is there anyone in the house?’ (Xu, 2007, p. 256)

With conditional contextual objRs, the referent exists only in the hypothetical world outlined in the protasis of a conditional construction (Croft, forthcoming, p. 85). There was one example of a CT Conditional Contextual ObjR, given in (58).

- (58) 恁 有 底 谁 欲 来 学习 ...  
 niŋ<sup>52</sup> u<sup>25-11</sup> ti<sup>11</sup>tian<sup>55</sup> ai<sup>21-53</sup> lai<sup>55-11</sup> hak<sup>5-2</sup>sip<sup>5</sup> ...  
 2PL have who want come learn ...  
 ‘[If] you (pl.) have anyone who wants to come study...’ [G1F6-24]

For indirect negation contextual objRs, the referent appears in a clause embedded in a negative clause (Croft, forthcoming, p. 85). An example of a Mandarin Indirect Negation

---

<sup>53</sup> An error in Xu’s (2007) transcription of this word has been corrected.

Contextual ObjRs is provided in (59). No examples were found in Cambodian Teochew or the literature.

- (59) 我 不 相信 任何 人 来 了  
 wǒ bù xiǎngxin rènhé rén lái le  
 I not think any man come PERF  
 ‘I don’t think that anyone came.’ (Haspelmath, 1997, p. 309)

With comparative contextual objRs, the unspecific referent appears in a comparative construction, as in the example in (60) from Mandarin.

- (60) 他 跑 得 比 任何 一个 男生 都 快  
 Tā pǎo de bǐ rènhé yīgè nánshēng dōu kuài  
 he run PT than any one boy indef fast  
 ‘He runs faster than any boy.’ (Haspelmath, 1997, p. 309)

No examples of this type were found in the current data set or the literature. Future research should explore the existence of these constructions in Cambodian Teochew (and other varieties).

With free choice contextual objRs, the identity of the unspecified referent can be freely selected without the changing proposition’s truth value (Croft, forthcoming, p. 86). There were no examples of these objRs in the current data set on Cambodian Teochew, though examples were found in Singapore Teochew, in (61) and Jieyang, in (62). More data should be collected to see if the construction is similarly found in Cambodian Teochew.

(61) tue<sup>53</sup> ku<sup>11</sup> li<sup>53</sup> si<sup>11</sup> mi<sup>ʔ5</sup> kai<sup>55</sup> loŋ<sup>33</sup> ɔi<sup>11</sup> hiou<sup>53</sup>  
 follow long 2SG what MOD all can  
 ‘After learning for a long time, you will be able to do anything.’  
 (Low, 2014, p. 21)

(62) luu<sup>53</sup> ti<sup>11</sup>tian<sup>33</sup>si<sup>55-11</sup> to<sup>33</sup> ho<sup>53-35</sup> lai<sup>55</sup>  
 2SG when all then come  
 ‘You can come anytime.’ (Xu, 2007, p. 250)

With the final type, direct negation contextual objRs, the unspecified referent is under the scope of the negation in the same clause (Croft, forthcoming, p. 86). Examples of CT Direct Negation Contextual ObjRs are given in (63). Note that (63b) additionally makes use of the universal quantifier 都 *tou<sup>33</sup>* ‘all’.

(63) a. 我 爸 在.. 就就 啊..  
 ua<sup>52</sup> pa<sup>33</sup> tɔ<sup>24</sup>.. tɕiu<sup>11</sup>, tɕiu<sup>11</sup> a<sup>33</sup>..  
 1SG.POSS father live.. then, then INTJ..  
 底噢 都 无 但 番家话 啊  
ti<sup>11</sup>tian<sup>55</sup> tou<sup>33</sup> bɔ<sup>11</sup> tã<sup>21-53</sup> huan<sup>33</sup>ke<sup>33-ue<sup>11</sup></sup> a<sup>33</sup>  
who all NEG speak local-language PRT  
 ‘When my father was alive, nobody spoke Khmer.’ [G1F6-24]

b. 无 有 底噢 来  
 bɔ<sup>55-11</sup> u<sup>24-11</sup> ti<sup>11</sup>tian<sup>55</sup> lai<sup>55</sup>  
 NEG have who come  
 ‘There is not anybody [who] comes.’ [G3F1-2]

#### 5.2.4 CT Generic ObjR Constructions

The last row in the information status continuum in Table 5.4 is for generic reference. In these constructions, the objR is referring to the type of objR and not any specific referent (Croft, forthcoming, p. 88).

Some CT Generic ObjR Constructions are shown in (64). Each of the underlined ObjRs in (64a) is referring to generic items of say the 鱼 *hu*<sup>55</sup> ‘fish’ type. There is not any specific ‘fish’ it is referring to. Instead, the speaker was extrapolating from the picture to tell his own story about the many delicious things to eat in Cambodia.

- (64) 有 物件 有 鱼 啊 有 肉  
 u<sup>11</sup> mueŋ<sup>55-2</sup>kiã<sup>24</sup> u<sup>11</sup> hu<sup>55</sup> a<sup>33</sup> u<sup>11</sup> baŋ<sup>2</sup>  
 have thing have fish PRT have meat  
 有 七个 菜.. 青果 七个 啊  
 u<sup>11</sup> miŋ<sup>55</sup>kai<sup>55</sup> tɕ<sup>h</sup>ai<sup>21</sup>.. tɕe<sup>33</sup>kuẽ<sup>52</sup> miŋ<sup>55</sup>kai<sup>55</sup> a<sup>33</sup>  
 have what vegetables.. fruit what PRT  
 ‘There’s things, there’s fish, there’s meat, there’s vegetables, fruit.’ [G2M1-5]

It has been proposed by Xu (2007, p. 108) that “bare” objR constructions (those with the objR with no classifier/modifier) in topic position in Jieyang are indicative of generic reference. Thus in (65) below, 猫 *ŋiau*<sup>33</sup> ‘cat’ can only be interpreted as referring to the type ‘cat’, and not any specific cat.

- (65) 猫 爱 食 鱼  
ŋiau<sup>33</sup> ai<sup>213-53</sup> tsiak<sup>5-2</sup> hu<sup>55</sup>  
cat love eat fish  
 ‘Cats like to eat fish.’ (Xu, 2007, p. 108)

In Cambodian Teochew, however, there is evidence that “bare” ObjRs can refer to a specific referent, as in (66). In (66a), which is the first sentence of the text, the speaker is referring to a specific cat, the main character in the forthcoming story. Similarly, (66b) is talking about specific children, the ones in the story. And in (66c), the 媵妘 *tɕa*<sup>33</sup>*bɔu*<sup>33</sup> ‘woman’

was referenced in the prior two utterances and thus is specific and cannot be interpreted generally.

- (66) a. 猫 死  
 ɲiau<sup>33</sup> i<sup>24</sup>  
 cat sleep  
 ‘[A] cat sleeps.’ [G1F1-9]
- b. 孛团 在 企 睇  
 nou<sup>33</sup>kiã<sup>52</sup> tɔ<sup>11</sup> k<sup>h</sup>ia<sup>24-11</sup> t<sup>h</sup>ɔ̃<sup>52</sup>  
 child PROG stand look  
 ‘[Some] children are standing looking.’ [G1F3-15]
- c. 媵妘 就 趁 伊 孛  
 tɕa<sup>33</sup>bɔu<sup>33</sup> tɕiu<sup>11</sup> dziau<sup>11</sup> i<sup>33</sup> tɕ<sup>h</sup>ɔk<sup>2</sup>  
 woman then chase away 3SG run.away  
 ‘[The] woman then chases him and he runs away.’ [G3F1-3]

There is some evidence that the pattern proposed by Xu (2007) does not hold in other Teochew varieties as well, such as in the following examples from Singapore Teochew. The example in (67a) follows an utterance about going to the hospital in a story. There is a specific ‘doctor’ who looked at the ‘nose’. The interpretation cannot be generic. Similarly, the ‘bicycle’ in (67b) is specific in interpretation as it is referring to the ‘bicycle’ in the story, which was also mentioned in the prior utterance.

- (67) a. lau<sup>33</sup>kuŋ<sup>33</sup> tã<sup>33</sup> ie<sup>31</sup> p<sup>h</sup>i<sup>33</sup> lian<sup>55</sup> tiŋ<sup>35</sup> k<sup>h</sup>i<sup>11</sup> si<sup>33</sup>tã<sup>33</sup>  
doctor say 3SG.POSS nose bridge break go but  
 bɔ<sup>11</sup> si<sup>33</sup>mi<sup>ʔ5</sup> tua<sup>11</sup> si<sup>11</sup>  
 NEG.have what big matter  
 ‘The doctor said his nose was broken but it wasn’t a big deal.’  
 (Low, 2014, p. 59)

- b.      lɔŋ<sup>2</sup>-lar<sup>11</sup>                      k<sup>h</sup>a<sup>33</sup>laŋ<sup>2</sup>ts<sup>h</sup>ra<sup>33</sup>    si<sup>11</sup>    paŋ<sup>53</sup>    tɔ<sup>11</sup>    t<sup>h</sup>ou<sup>11</sup>    ɛ<sup>35</sup>    kɔ<sup>11</sup>  
 descend-COMP.DIR bicycle                      be put    LOC ground below there  
 ‘(When he) came down...the bicycle is put on the ground below.’  
 (Yeo, 2011, p. 98)

Matthews and Pacioni (1997) looked at the contrast between this generic and specific interpretation of “bare” objRs in Cantonese and Mandarin. In Cantonese, “bare” objRs have generic interpretation, while Mandarin allows a specific reading for such objRs. Matthews and Pacioni suggest that “bare” objRs with specific reference in Cantonese can potentially occur if they have a “name-like function” within a story/fable (Matthews & Pacioni, 1997, p. 50). This could be the case for CT examples like (68) below, where the main protagonist of the story, the 猫 *ŋiau*<sup>33</sup> ‘cat’, could be referenced by a name like ‘Cat’.

- (68) a.      □      猫      行      在      路  
           u<sup>33</sup>      ŋiau<sup>33</sup>      kia<sup>55-11</sup>      tɔ<sup>11</sup>      lou<sup>11</sup>  
           INTJ    cat      walk    LOC    road  
           ‘Cat walks on [the] road.’ [G1F1-9]
- b.      猫                      走                      了  
           ŋiau<sup>33</sup>                      tɕau<sup>52-24</sup>                      liau<sup>52</sup>  
           cat                      run                      already  
           ‘Cat has gone (across the river).’ [G1F1-9]

However, this explanation does not seem to hold for all of the instances where specific bare ObjRs occur in Cambodian Teochew, like (66b) and (66c) for example, where the referents 孛囡 *nɔu*<sup>33</sup>*kiã*<sup>52</sup> ‘child’ and 媪妯 *tɕa*<sup>33</sup>*bɔu*<sup>33</sup> ‘woman’ are not likely to have a name-like function since they are only peripheral participants. Therefore, I propose that the interpretation of CT “Bare” ObjRs is context-dependent, rather than being strictly generic or strictly specific.

Given the examples in (67), this may be the case for Singapore Teochew as well, though more examples are needed from that variety. Simpson, Soh, and Nomoto (2011, p. 184) claim that many Cantonese speakers also accept the use of “bare” objRs for specific reference. As Mandarin allows a specific reading for such bare objRs, there is potential that these languages are showing an influence from Mandarin.

As for Cambodian Teochew, since classifiers are optional in Khmer, there is no requirement as to an objR having a classifier in order to lend a specific interpretation. Thus the “bare” ObjR in (69) does not have a generic interpretation. There is potential that this usage of “bare” ObjRs in Khmer has influenced their usage in Cambodian Teochew.

- (69) ឆ្កែ ខាំ កាត់  
ckae k<sup>h</sup>am koat  
dog bite 3SG  
'The dog bit him.' (Sak-Humphry, 2016, p. 221)

### 5.3 Extended Animacy Hierarchy

Croft (forthcoming, p. 61) outlines common semantic classes for prototypical referents, or objRs:

1. **Humans:** 大夫 *ta<sup>33</sup>pəu<sup>33</sup>* ‘man’, 媯妘 *tɕa<sup>33</sup>bəu<sup>52</sup>* ‘woman’, 媯妘囡 *tɕa<sup>33</sup>bəu<sup>52</sup>kiã<sup>52</sup>* ‘girl’
  - a. kinship: 爸爸 *pa<sup>55-11</sup>pa<sup>55</sup>* ‘father’, 弟弟 *ti<sup>24-11</sup>ti<sup>24</sup>* ‘younger brother’, 姨 *i<sup>55</sup>* ‘aunt’
  - b. social role: 马打 *ma<sup>33</sup>ta<sup>24</sup>* ‘police’, 厝边人 *tɕ<sup>h</sup>u<sup>21-53</sup>pi<sup>33</sup>naŋ<sup>55</sup>* ‘neighbor’
2. **Animals:** 狗 *kau<sup>52</sup>* ‘dog’, 牛 *gu<sup>55</sup>* ‘cow’, 鸟 *tɕiau<sup>52</sup>* ‘bird’
3. **Plants:** 树 *tɕ<sup>h</sup>ui<sup>11</sup>* ‘tree’, 花 *hue<sup>33</sup>* ‘flower’, 草 *tɕ<sup>h</sup>au<sup>52</sup>* ‘grass’
4. **Animal and plant products:** 苹果 *p<sup>h</sup>ej<sup>11</sup>kuẽ<sup>52</sup>* ‘apple’, 肉 *baŋ<sup>2</sup>* ‘meat’, 菜 *tɕ<sup>h</sup>aŋ<sup>21</sup>* ‘vegetables’
5. **Artifacts:** 衫 *sã<sup>33</sup>* ‘shirt’, 索 *sɔŋ<sup>2</sup>* ‘rope’, 电脑 *tiaŋ<sup>55-11</sup>nau<sup>52</sup>* ‘computer’
6. **Body parts (human, animal, plant):** 脚 *k<sup>h</sup>a<sup>33</sup>* ‘foot’, 头 *t<sup>h</sup>au<sup>55</sup>* ‘head’, 尾 *bue<sup>52</sup>* ‘tail’, 角 *kak<sup>2</sup>* ‘horn’, 叶 *hiɔŋ<sup>5</sup>* ‘leaf’
7. **Natural (inanimate) object/substances:** 水 *tɕui<sup>52</sup>* ‘water’, 火 *hue<sup>52</sup>* ‘fire’, 涂 *t<sup>h</sup>əu<sup>55</sup>* ‘dirt’
8. **Places:** 山 *suã<sup>33</sup>* ‘mountain’, 海 *har<sup>52</sup>* ‘sea’, 岛 *tau<sup>52</sup>* ‘island’, 厝 *tɕ<sup>h</sup>u<sup>21</sup>* ‘house’, 市 *tɕ<sup>h</sup>i<sup>24</sup>* ‘city’, 花园 *hue<sup>33</sup>huŋ<sup>55</sup>* ‘garden’

The eight classes are typically divided into three semantic groups of human (class 1), (nonhuman) animate (class 2), and inanimate (classes 3-8) (Croft, forthcoming, p. 61). These semantic categories of objRs make up the Extended Animacy Hierarchy (EAH) in Figure 5.1, adapted from Croft (2003, p. 130), which is also called ‘referentiality hierarchy’.

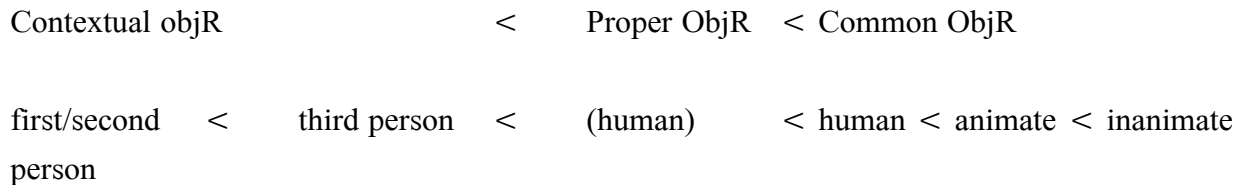


Figure 5.1: Extended Animacy Hierarchy

Languages frequently make distinctions based on the EAH, some of which were first noted in Silverstein (1976) and Dixon (1979) in regards to ergativity. But there are many other ways that distinctions in language are made according to the EAH. This section highlights one such distinction for Cambodian Teochew.

One of the common distinctions based on the EAH is in regards to grammatical number. CT Personal Contextual ObjR Constructions are no exception. As was shown in §5.1.3.1, while

the more salient CT First and Second Person Contextual ObjRs have plural forms that are one syllable, the Third Person Plural has two syllables (and two characters). Furthermore, it was shown that at least for some speakers in some contexts there is no singular/plural distinction in Third Person, with 伊 *i*<sup>33</sup> being used for both. The fact that the plural distinction exists at all in CT Personal Contextual ObjRs is further support of the EAH. Plural distinctions are not present on ObjRs further down the hierarchy, including for humans (70a), animals (70b), and objects (70c). The form of the ObjR is identical in each pair no matter if it is singular or plural.

(70)	a.	人	二	人
		naŋ <sup>55</sup>	nɔ <sup>24-11</sup>	naŋ <sup>55</sup>
		person	two	person
		‘[a] person’		‘two people’
	b.	蜀 隻 狗	四 隻 狗	
		tʃek <sup>5-2</sup> tʃiaŋ <sup>2-5</sup> kau <sup>52</sup>	si <sup>21-53</sup> tʃiaŋ <sup>2-5</sup> kau <sup>52</sup>	
		one CLF dog	four CLF dog	
		‘one dog’		‘four dogs’
	c.	簿	三	本 簿
		p <sup>h</sup> ɔu <sup>24</sup>	sã <sup>33</sup>	puŋ <sup>52-24</sup> p <sup>h</sup> ɔu <sup>24</sup>
		book	three	CLF book
		‘[a] book’		‘three books’

#### 5.4 CT Property Reference Constructions

Property reference constructions exhibit reference to a property. According to Croft (forthcoming), these constructions are “very rare in discourse, little described in reference grammars, and little studied in typology” (p. 599), so they are not extensively covered in his forthcoming book. Overall there were not a lot of property words in general in my collected data and thus there was little evidence on the existence or behavior of CT Property Reference Constructions.

In the literature on Teochew, there are some mentions of property reference constructions. Xu (2007, pp. 66-7) provided examples of pairs of property words that combine to form Jieyang Common ObjR Constructions, shown in (71).

- |      |    |                                       |    |   |    |   |
|------|----|---------------------------------------|----|---|----|---|
| (71) | a. | tua <sup>11</sup> -soi <sup>213</sup> | b. | ŋaŋ <sup>55-11</sup> -zua? <sup>5</sup> | c. | pui <sup>55-11</sup> -saŋ <sup>53</sup>   |
|      |    | big-small                             |    | cold-hot                                |    | fat-thin                                  |
|      |    | ‘size’                                |    | ‘weather conditions’                    |    | ‘body build’                              |
|      | d. | ke <sup>33</sup> -kiam <sup>53</sup>  | e. | kuĩ <sup>55-11</sup> -ke <sup>35</sup>  | f. | kiam <sup>55-11</sup> -tsiã <sup>53</sup> |
|      |    | more-less                             |    | high-low                                |    | salty-bland                               |
|      |    | ‘total amount’                        |    | ‘height’                                |    | ‘(taste of a) dish’                       |

Xu (2007, p. 43) also provided examples of words formed with an affix 性 *sẽ<sup>213</sup>* ‘-ity, -ness’, in (72), which she claimed has been borrowed from Mandarin.

- |      |    |  |    |   |
|------|----|--|----|---|
| (72) | a. | pek <sup>2</sup> ziaŋ <sup>35-21</sup> - <u>sẽ<sup>213</sup></u> | b. | ŋiam <sup>55</sup> toŋ <sup>35-21</sup> - <u>sẽ<sup>213</sup></u> |
|      |    | ‘ <u>necessity</u> ’   |    | ‘ <u>seriousness</u> ’ (Xu, 2007, p. 43)                          |

No evidence was found for the constructions in (71) and (72) in Cambodian Teochew, but consultants were not specifically asked about them. Future research could inquire about them.

### 5.5 CT Action Reference Constructions

Actions can act as arguments of a predicate. In these cases, actions are performing the information packaging function of reference. These are called action reference constructions.

Croft (forthcoming, pp. 42-43) presented examples of English Action Reference Constructions, along with their terminology in traditional grammars in (73).

- (73) a. The explosion startled them. [Nominalization]  
 b. Hiking in Canyonlands is challenging. [Gerund]  
 c. They want to eat in the kitchen. [Infinitival Complement]  
 d. Frieda thinks (that) Janet won't come to the party. [Finite Complement]

While traditional grammars often make a sharp distinction between the (a) and (b) examples compared to the (c) and (d) examples, in RCG all are considered action reference constructions.<sup>54</sup>

Not all predicates can take actions as arguments. Croft (forthcoming) outlines several different semantic types of action-taking predicates or ATPs<sup>55</sup> (pp. 513-520), following Noonan (2007). Here I explore some of the ones found in my data on Cambodian Teochew, after a brief look of the types of strategies used in action reference constructions.

CT Action Reference Constructions use a balanced strategy (from Croft, forthcoming, pp. 442-3) where the predicate form is the same as what it would look like in a simple declarative main clause. Thus all of the underlined portions of the examples that follow in (74-79, 81) can be uttered on their own with no change to their form. We can also look at action reference constructions in language according to the strategy or strategies used for argument structure. In constructions using a split argument structure strategy, there are two clauses and two distinct argument structure constructions (Croft, forthcoming, p. 531). Conversely, constructions with only a single argument structure are said to use a merged argument structure strategy (Croft, forthcoming, p. 531). In a partially-merged argument structure strategy, the two clauses in the construction have one or more shared participants (Croft, forthcoming, p. 531).

---

<sup>54</sup> For Croft (forthcoming), these are all called “complement clause constructions”.

<sup>55</sup> Croft (forthcoming, p. 512) uses the terminology from Noonan (2007, p. 53), calling these “complement-taking predicates” or CTPs.

We highlight below the potential analyses in regards to the argument structure strategies used in CT Action Reference Constructions.

CT Action Reference Constructions can be used for utterance events. These constructions can use the direct report strategy in (74), where the speech is reported in the same form as it was said. Alternatively, they can use the indirect report strategy which only presents the content of the utterance, as in (75). In (75), the use of the indirect report strategy is clear because of the Third Person Contextual ObjRs (rather than First Person). These examples all use the split argument structure strategy.

- (74) a. 我 咀 “拜拜”  
 ua<sup>52</sup> tã<sup>21</sup> “pai<sup>33</sup>pai<sup>33</sup>”  
 1SG say “byebye”  
 ‘I say “byebye”.’ [G2M1-5]
- b. 汝 咀 “Ø 欢喜”  
 lu<sup>52</sup> tã<sup>21</sup> “Ø huã<sup>33</sup>hi<sup>52</sup>”  
 2SG say “1SG happy”  
 ‘You say “[I am] happy”.’ [G2M1-5]
- c. 我 问 汝 “欢喜 套 来 高棉?”  
 ua<sup>52</sup> muŋ<sup>11</sup> lu<sup>52</sup> “huã<sup>33</sup>hi<sup>52</sup> bɔi<sup>24</sup> lai<sup>55-11</sup> kau<sup>33</sup>mian<sup>55?</sup>”  
 1SG ask 2SG happy NEG.can come Cambodia  
 ‘I ask you “Are [you] happy to come to Cambodia?”’ [G2M1-5]

- (75) a. 只 个 嫲妸 合 人 .. 合 只 个  
 tɕi<sup>52</sup> kai<sup>55-11</sup> tɕa<sup>33</sup>bou<sup>52</sup> kaʔ<sup>2</sup> naŋ<sup>55-11</sup> .. kaʔ<sup>2</sup> tɕi<sup>52</sup> kai<sup>55-11</sup>  
 this LW woman with people .. with this LW  
 马打 但 但 伊 个 翁人 拍 伊  
 ma<sup>33</sup>ta<sup>24</sup> tā<sup>21</sup> tā<sup>21-53</sup> i<sup>33</sup> kai<sup>55-11</sup> aŋ<sup>33</sup>naŋ<sup>55</sup> paʔ<sup>2-5</sup> i<sup>33</sup>  
 police say COMP 3SG LW husband hit 3SG  
 ‘This woman told people.. told these police officers that her husband hit her.’ [G3F2-4]
- b. 但 伊 伊 无 好 □  
 tā<sup>21-53</sup> i<sup>33</sup>, i<sup>33</sup> bə<sup>55-11</sup> haũ<sup>21</sup> nia<sup>33</sup>  
 say 3SG, 3SG NEG willing PRT  
 ‘He says he is not willing [to drink alcohol].’ [G3F2-4]

CT Action Reference Constructions can be used for propositional attitude events such as thinking (76a), knowing (76b), and knowing how to do something (76c). The first sentence in (76a) shows how the ATP ‘think’ does not require an Action Reference Construction, as it requires no argument at all.

- (76) a. 猫 伊 想 □. 伊 想 食 鱼  
 ŋiau<sup>33</sup> i<sup>33</sup> sið<sup>24</sup> ne<sup>55</sup>. i<sup>33</sup> sið<sup>24-11</sup> tɕiaʔ<sup>5-2</sup> hu<sup>55</sup>  
 cat 3SG think PRT. 3SG think eat fish  
 ‘The cat, he thinks. He thinks of eating fish.’ [G1F2-14]
- b. Ø 唔 知 伊 嫲 在 块 □  
 Ø m<sup>11</sup> tɕai<sup>33</sup> i<sup>33</sup> ai<sup>55</sup> tɔ<sup>11</sup> kɔ<sup>21</sup> nia<sup>11</sup>  
 1SG NEG know 3SG mother LOC where PRT  
 ‘[I] don’t know where his mother is.’ [G2M1-6]
- c. 我 唔 晓 但 只个  
 ua<sup>52</sup> m<sup>11</sup> hiau<sup>24</sup> tā<sup>21</sup> tɕi<sup>52</sup>kai<sup>55</sup>  
 1SG NEG know.how say PDCO  
 ‘I don’t know how to say this.’ [G2M1-5]

CT Action Reference Constructions can be used for non-factive situations such as fearing events (77) and hoping events (78). These examples both use the split argument structure strategy.

(77) 伊 惊 我 食 伊 啊  
 i<sup>33</sup> kiã<sup>33</sup> ua<sup>52</sup> tɕia?<sup>5-2</sup> i<sup>33</sup> a<sup>33</sup>  
 3SG fear 1SG eat 3SG PRT  
 ‘He is afraid I [will] eat him.’ [G2F1US-18]

(78) 自早 希望 但 ...Ø 带动 人 只块 ...  
 tɕu<sup>11</sup>tɕa<sup>52</sup> hi<sup>33</sup>mɔ<sup>11</sup> tã<sup>21-53</sup> ...Ø tua<sup>53</sup>tɔŋ<sup>24-11</sup> nan<sup>55</sup> tɕi<sup>52</sup>kɔ<sup>21</sup> ...  
 just now hope COMP ...[drama troupe]inspire people here ...  
 ‘Since the beginning I have hoped that [the Teochew drama troupe] will inspire the people here.’ [G1F7-24]

CT Action Reference Constructions can be used for perception events such as seeing, as in (79). In each example, there is the shared participant, where the person in the first clause sees someone or something, and in the second clause, that seen someone is doing something.

- (79) a.  $\emptyset$  睇 伊 团 兜  
 $\emptyset$  t<sup>h</sup>ɔ̃i<sup>52-24</sup> i<sup>33</sup> kiã<sup>52</sup> i<sup>24</sup>  
 3SG see 3SG.POSS child sleep  
 ‘[She] sees her child sleeping.’ [G2M1-6]
- b. 我 睇见 只 个 口 媵妘 伊  
 ua<sup>52</sup> t<sup>h</sup>ɔ̃i<sup>52</sup> kian<sup>21</sup> tɛi<sup>52</sup> kai<sup>55-11</sup> u tɛa<sup>33</sup> bɔu<sup>52</sup> i<sup>33</sup>  
 1SG see this LW INTJ woman 3SG  
 在 但 合 只 个 大夫  
 tɔ<sup>11</sup> tã<sup>21-53</sup> kaɿ<sup>2-5</sup> tɛi<sup>52</sup> kai<sup>55-11</sup> ta<sup>33</sup> pou<sup>33</sup>  
 PROG talk with this LW man  
 ‘I see this woman talking to this man.’ [G3F2-4]
- c. ...就  $\emptyset$  睇见 鱼 泅 经过 伊  
 ...tɛiu<sup>11</sup>  $\emptyset$  t<sup>h</sup>ɔ̃i<sup>52-11</sup> kian<sup>21-53</sup> hu<sup>55</sup> siu<sup>55-11</sup> ken<sup>33</sup> kue<sup>21-53</sup> i<sup>33</sup>  
 ...then 3SG see fish swim pass by 3SG  
 ‘...then he sees fish swimming past him.’ [G3F1-3]

The examples in (80) showcase the difference between an English Action Reference Construction with a perception predicate using a split argument structure strategy (80a) and a partially-merged argument structure strategy (80b). In (80a), the entire clause [he is walking to the store] is the argument of the predicate *see*. Conversely, in (80b), *him* is a participant of the predicate *see* and at the same time is a participant of the predicate *walking*. (80a) also uses a balanced strategy where the predicate *is* has the same form as it would in a simple declarative main clause. (80b) uses a deranked strategy where the predicate differs from a declarative main clause predicate (Croft, forthcoming, p. 358). The English examples are able to be analyzed in such a way because of the difference in the form of the predicate (balanced vs. deranked), as well as the difference in the case of *he/him*. These same distinctions are not made in Cambodian Teochew and thus it is more difficult to determine if the examples in (79) are using a split argument structure strategy or a partially-merged argument structure strategy.

- (80) a. I see [he is walking to the store].  
 b. I see him walking to the store.

CT Action Reference Constructions can be used for desiderative events that express wanting, as in (81). (81a-b) have a shared participant in 伊 *i*<sup>33</sup> ‘3SG’ and are using either a partially merged or potentially fully merged argument structure strategy. Similar to the examples with perceptive predicates in (79), (81c) could be analyzed as using either a split argument structure or a partially merged argument structure strategy.

- (81) a. 伊 欲 哭  
*i*<sup>33</sup> *aĩ*<sup>21-53</sup> *k*<sup>h</sup>*au*<sup>21</sup>  
 3SG want cry  
 ‘He wants to cry.’ [G2M1-6]
- b. □ 伊 欲.. 觅 头家 食 物件  
*ɔ*<sup>21</sup> *i*<sup>33</sup> *aĩ*<sup>21</sup>.. *tɕ*<sup>h</sup>*ue*<sup>11</sup> *t*<sup>h</sup>*au*<sup>55-11</sup> *ke*<sup>33</sup> *tɕia*<sup>ʔ5-2</sup> *mue*<sup>ʔ5-2</sup> *kiã*<sup>24</sup>  
 INTJ 3SG want.. find store.owner eat thing  
 ‘It wants to find its owner and eat things.’ [G1F1-9]
- c. Ø 欲 主人 互 饭 食  
 Ø *aĩ*<sup>21</sup> *tɕu*<sup>24</sup> *jin*<sup>55</sup> *hɔu*<sup>11</sup> *puŋ*<sup>11</sup> *tɕia*<sup>ʔ5</sup>  
 [3SG] want master give rice eat  
 ‘[It] wants [its] master to give [it] rice to eat.’ [G1F1-9]

CT Action Reference Constructions can optionally use the complementizer<sup>56</sup> strategy where a certain morpheme appears between the ATP and the referred action. In Cambodian Teochew, the CT Action Predicate 𠵼 *tã*<sup>21</sup> ‘say, speak, tell’ has grammaticalized into a complementizer, as it has in Jieyang (Xu, 2007, p. 141). Chappell (2008) presented five stages

---

<sup>56</sup> I have chosen to keep the term complementizer strategy from Croft (forthcoming, p. 520) rather than coin a new term. This is reflected in the glossing of CT Action Reference Constructions.

of grammaticalization of ‘say’ verbs to complementizer based on data from ten Sinitic languages, including Taiwanese Southern Min, a language closely related to Teochew. We show her stages for Cambodian Teochew below. Before it grammaticalized, there are the lexical usages of the CT Action Predicate 𠵼 *tã<sup>21</sup>* ‘say, speak, tell’ which can be quite varied, as shown in (82).

- (82) a. 好 伊 妈妈.. 𠵼 故事  
*hɔ<sup>52</sup> i<sup>33</sup> ma<sup>33</sup>ma<sup>33</sup> .. tã<sup>21-53</sup> ku<sup>53</sup>su<sup>11</sup>*  
 good 3SG.POSS mama.. tell story  
 ‘Then his mother tells a story.’ [G1F1-10]
- b. Ø 坐 在 𠵼话  
 Ø *tɕɔ<sup>11</sup> tɔ<sup>11</sup> tã<sup>21-53</sup>ue<sup>11</sup>*  
 3PL sit PROG say-word  
 ‘They sit talking.’ [G1F1-10]
- c. ... 在 𠵼 国语  
 ... *tɔ<sup>11</sup> tã<sup>21-53</sup> kɔk<sup>2-5</sup>gu<sup>52</sup>*  
 ... PROG speak Mandarin  
 ‘...[you] are speaking Mandarin.’ [G1F6-24]

In Chappell’s (2008, p. 59) Stage 1 of grammaticalization, ‘say’ is used as quotative, where it introduces either direct (83a) or indirect (83b) speech.

- (83) a. 伊 𠵼 “□! 真 大”  
*i<sup>33</sup> tã<sup>21</sup> “ua<sup>33</sup>! tɕiŋ<sup>33</sup> tua<sup>11</sup>*  
 3SG say EXLA very big  
 ‘He says, “Wow! So big!”’ [G2F1US-20]
- b. 伊 𠵼 伊 无 食 □  
*i<sup>33</sup> tã<sup>21-53</sup> i<sup>33</sup> bɔ<sup>55-11</sup> tɕiaŋ<sup>5</sup> nia<sup>33</sup>*  
 3SG say 3SG NEG eat PRT  
 ‘He says he doesn’t drink [alcohol].’ [G3F2-4]

In Chappell's (2008, p. 59) Stage 2 a speech act predicate such as 'ask' or 'tell' combines with 'say', as in (84). Chappell calls this a "serial verb construction" but here we'll say it uses a serial predicate strategy (see §7.1.4.1).

- (84) ... 伊 问 但 “日日日日 我 ...”  
 ... i<sup>33</sup> muŋ<sup>11</sup> tā<sup>21</sup> “dzik<sup>5-2</sup>dzik<sup>5</sup>dzik<sup>5-2</sup>dzik<sup>5</sup> ua<sup>52</sup> ...”  
 ... 3SG ask say “day RED day RED 1SG ...”  
 ‘...he asks saying “Every day I...”’ [G2F1US-19]

Chappell makes a distinction between examples like those in (84) and those in (85). When there is an intervening referent between the speech act predicate and the 'say' predicate, 'say' is non-grammaticalized (Chappell, 2008, p. 71). The interpretation of the examples (85) is that of 'ask X saying' while those in (84) are interpreted either as 'ask saying' or 'ask that'.

- (85) a. ...伊 问 阿妈 但 “阿妈 俺 怎呢...”  
 ...i<sup>33</sup> muŋ<sup>11</sup> a<sup>33</sup>ma<sup>55</sup> tā<sup>21-53</sup> “a<sup>33</sup>ma<sup>55</sup> naŋ<sup>52</sup> tɕɔ<sup>52</sup>ni<sup>55</sup>...”  
 ...3SG ask mother say mom 1PL why...  
 ‘... he asks his mother saying, “Mom, why..”’ [G2F1US-17]
- b. 伊人 问 伊 但 怎呢 ...  
 i<sup>33</sup>- naŋ<sup>55</sup> muŋ<sup>11</sup> i<sup>33</sup> tā<sup>21-53</sup> tɕɔ<sup>52</sup>ni<sup>55</sup> ...  
 3SG- person ask 3SG say why ...  
 ‘People ask him saying, why...’ [G3F2-4]

In Chappell's (2008) Stage 3, 'say' is used as a complementizer with cognitive predicates like 'think' and 'know', shown in (86). Unlike with the possible intervening referent with predicates in Stage 2 in (85), there is never a break between the cognitive predicate and the 'say' predicate; they are a tightly bonded unit (Chappell, 2008, p. 61).

- (86) a. Ø 在 想 但 先回 伊 在 好好  
 Ø tɔ<sup>11</sup> sĩ<sup>24-11</sup> tã<sup>21-53</sup> sai<sup>33</sup> hue<sup>55</sup> i<sup>33</sup> tɔ<sup>11</sup> hɔ<sup>52-11</sup> hɔ<sup>52</sup>  
 3SG PROG think COMP before 3SG PROG good RED  
 ‘He is thinking that before he was good.’ [G3F2-4]
- b. 伊..伊 就 想 但.. 就.. Ø 欲 做 好 人  
 i<sup>33</sup>..i<sup>33</sup> tɕiu<sup>11</sup> sĩ<sup>24-11</sup> tã<sup>21-53</sup>.. tɕiu<sup>11</sup>.. Ø ai<sup>21-53</sup> tɕɔ<sup>52</sup> hɔ<sup>52</sup> naŋ<sup>55</sup>  
 3SG then think COMP.. then.. 3SG want do good person  
 ‘Then he thinks that he wants to be a good person.’ [G3F3-4]
- c. 只个 Ø 知 但 国语 木瓜 ...  
 tɕai<sup>55</sup> Ø tɕai<sup>33</sup> tã<sup>21-53</sup> kɔk<sup>2-5</sup> gu<sup>52</sup> mùguā ...  
 PDCO 1SG know COMP Mandarin papaya ...  
 ‘This (one), I know that Mandarin [is] mùguā ...’ [G1F5US-23]

In Chappell’s (2008) Stage 4, more predicates are able to combine with complementizer ‘say’ including those of speech acts, cognition and perception, and emotion. In (87), we see ‘say’ used with hoping events. Like those predicates in Stage 3, there can be no intervening referent in between the first predicate and the ‘say’. There is also no pause between those two words. In (86b) and (87b), the speakers do take a pause, as indicated by the two dots (..), but this pause happens after the predicate-complementizer compound, as the two are “now a virtually indivisible unit” (Chappell, 2008, p. 61).

- (87) a. 我 希望 但 有蜀日 口 再 有  
 ua<sup>52</sup> hi<sup>33</sup> mo<sup>11</sup> tã<sup>21-53</sup> u<sup>11</sup> tcek<sup>5-2</sup> dzik<sup>5</sup> na<sup>11</sup> tcai<sup>21-53</sup> u<sup>11</sup>  
 1SG hope COMP one day LOC again have  
 只 个 潮州.. 俺 潮州 个 潮剧- 俺 个  
 tci<sup>52</sup> kai<sup>55-11</sup> tie<sup>11</sup> tciu<sup>33</sup> .. nan<sup>52</sup> tie<sup>11</sup> tciu<sup>33</sup> kai<sup>55-11</sup> tie<sup>11</sup>-kia<sup>ʔ5</sup>- nan<sup>52</sup> kai<sup>55-11</sup>  
 this LW Teochew.. 1PL.POSS Teochew LW TC-drama- 1PL LW  
 潮 潮剧 团 来 只块 来,来 只块  
 tio<sup>11</sup>- tio<sup>11</sup>-kia<sup>ʔ5-2</sup> t<sup>h</sup>uan<sup>55</sup> lai<sup>55-11</sup> tci<sup>52</sup> ko<sup>21</sup> lai<sup>55</sup>, lai<sup>55</sup> tci<sup>52</sup> ko<sup>21</sup>  
 TC- TC-drama troupe come here come, come here  
 ‘I hope that one day our Teochew drama troupe comes here again.’  
 [G1F7-24]
- b. 自 早 Ø 希望 但.. Ø 带动 人 只块 ...  
 tciu<sup>11</sup> tca<sup>52</sup> Ø hi<sup>33</sup> mo<sup>11</sup> tã<sup>21-53</sup> .. Ø tua<sup>53</sup> toŋ<sup>24-11</sup> nan<sup>55</sup> tci<sup>52</sup> ko<sup>21</sup> ...  
 since early 1SG hope COMP.. 3SG inspire people here ...  
 ‘Since the beginning I have hoped that [the Teochew drama troupe] will  
 inspire the people here...’ [G1F7-24]

In Chappell’s (2008, p. 62) Stage 5, the usage of ‘say’ as a complementizer has expanded into modal predicates like ‘must’. There were no examples of this usage in the current data set. However, Xu and Matthews (2007, p. 65) provided evidence that Chaozhou dialects (as well as Taiwanese Southern Min) have reached this stage of grammaticalization. More data will need to be gathered to see if this is the case in Cambodian Teochew as well.

Overall, evidence points to ‘say’ being grammaticalized as a complementizer in Cambodian Teochew at least as far as Chappell’s Stage 4. And it exhibits other characterizations of its grammaticalization as noted by Chappell (2008, p. 62) for Sinitic languages: ‘say’ retains its lexical usages, as shown in (82), and is optional as a complementizer, as evidenced by the examples of CT Action Reference Constructions that do not use the complementizer strategy such as (76) and (77). Finally, the complementizer ‘say’

can be used with itself, as in (88), which suggests a “high degree of conventionalization” (Chappell, 2008, p. 62).

- (88) 只 个 嫫妘 ... 但 但 伊 个 翁人 拍 伊  
 tɕi<sup>52</sup> kai<sup>55-11</sup> tɕa<sup>33</sup>bou<sup>52</sup> ... tã<sup>21</sup> tã<sup>21-53</sup> i<sup>33</sup> kai<sup>55-11</sup> aŋ<sup>33</sup>naŋ<sup>55</sup> paɿ<sup>2-5</sup> i<sup>33</sup>  
 this LW woman ... say COMP 3SG LW husband hit 3SG  
 ‘This woman ... said that her husband hit her.’ [G3F2-4]

## 5.6 Summary

This chapter explored varying constructions that express the information packaging function of reference, including CT Object Reference Constructions, property reference constructions, and CT Action Reference Constructions. We also covered different constructions that indicate certain statuses of referents based on the information status continuum, as well as distinctions made in Cambodian Teochew that align with the Extended Animacy Hierarchy. Throughout the chapter we have shown some of the ways that Cambodian Teochew is similar to other Teochew varieties, and have also highlighted some of the differences, which include: fewer First Person Plural Contextual ObjR Constructions, at least for some speakers; the use of Third Person 伊 *i*<sup>33</sup> for both singular and plural contexts; and “bare” ObjRs whose identity is known (contrary to that reported in Jieyang).

Ultimately we have merely scratched the surface in exploring CT Reference Constructions and as such throughout the chapter we have also highlighted many areas that would benefit from future research.

## Chapter 6 – Modifiers and (Referent) Modification Constructions

This chapter explores CT Modification Constructions. Modification can be described as a “speaker adds further information about the referent” (Croft, forthcoming, p. 93). Referent modification constructions, or just modification constructions, consist of a modifier and a referring expression, such as the CT Modification Construction 老车 *lau<sup>24-11</sup> tɕ<sup>h</sup>ia<sup>33</sup>* ‘old car’. From Croft (forthcoming, p. 93), the terms modifier, modification, and attributive phrase refer to the modifier of referring expressions, such as the 老 *lau<sup>24-11</sup>* ‘old’ in the prior example. Croft (forthcoming, p. 94, citing Croft 1990, 2007) identified three subfunctions of the information packaging function of modification: subcategorizing or sorting, situating, and selecting.

Subcategorizing limits the interpretation of the type of the referent. For example, the CT Common ObjR 马 *be<sup>52</sup>* ‘horse’ refers to all items that are of the ‘horse’-type. But the CT Modification Construction 乌马 *ɔu<sup>33</sup> be<sup>52</sup>* ‘black horse’ can only refer to those ‘horse’-type items which are ‘black’. Thus the property concept, the CT Modifier Construction 乌 *ɔu<sup>33</sup>* ‘black’, is subcategorizing the referent *be<sup>52</sup>* ‘horse’. Property concepts, which are the prototypical modifiers, perform this subfunction. CT Property Modification Constructions are explored in §6.1.

Selecting highlights the intended instance(s) or set of instances of a referent (Croft, forthcoming, p. 98). Cardinal enumeration is one example of the selecting subfunction. In the CT Enumeration Modification Construction 六个球 *lak<sup>5-2</sup> kai<sup>55-11</sup> kiu<sup>55</sup>* ‘six balls’, the CT Cardinal Enumeration Construction 六 *lak<sup>5-2</sup>* ‘six’ selects six and only six instances of the referent 球 *kiu<sup>55</sup>* ‘ball’. Modification constructions that perform the subfunction of selecting are covered in §6.2.

Situating more specifically identifies a referent with additional information such as its location or relationship to others (Croft, forthcoming, p. 94). Chapter 5 of this dissertation explored how deixis and information status can be used to situate a referent. For example, in the CT Deictic Contextual Attribute Construction 只个路  $t\check{c}i^{52} ka^{55-11} lu^{11}$  ‘this road’, the speaker can only be referring to a particular ‘road’ that has proximal deixis, and thus is relatively nearby. Similarly, in the CT Anaphoric Article Constructions 伊囡  $i^{33} ki\tilde{a}^{52}$  ‘the child’ and 隻狗  $t\check{c}ia^{2-5} kau^{52}$  ‘the dog’ in the contexts they are spoken by G1F1 in text #11, the modifiers  $i^{33}$  and  $t\check{c}ia^{2-5}$  situate the referents  $ki\tilde{a}^{52}$  and  $kau^{52}$  respectively where they can only be understood to be referring to one particular ‘child’ and one particular ‘dog’. Object concepts can also be used in this manner. In the CT Common ObjR Construction 妈妈  $ma^{55-11} ma^{55}$  ‘mother’,  $ma^{55-11} ma^{55}$  can refer to all items which are the ‘mother’-type, while the CT Modification Construction  $ua^{52} ma^{55-11} ma^{55}$  ‘my mother’ as uttered by me can only refer to one entity, the ‘mother’-type person that relates to me. The object concept,  $ua^{52}$  ‘1SG’, is modifying the referent,  $ma^{55-11} ma^{55}$  ‘mother’.

Object concepts can also subcategorize referents. While the CT Common ObjR Construction 肉  $ba^{2}$  ‘meat’ refers to all things ‘meat’-type, the CT Modification Construction 羊肉  $i\tilde{\sigma}^{55-11} ba^{2}$  literally ‘sheep-meat’, interpreted as ‘lamb, mutton’, refers to only ‘meat’-type items that are of the ‘sheep’ subcategory. CT Object Modification Constructions, of both the subcategorizing and situating subfunctions, are explored in §6.3.

Finally, action concepts can perform the information packaging function of modification. CT Action Modification Constructions are introduced in §6.4.

## 6.1 CT Property Modification Constructions

Property modification constructions<sup>57</sup> sort or subcategorize referents. Property modifiers can be divided into subtypes according to Table 6.1. Examples of CT Property Modification Constructions are provided in the right column.

Information packaging function	Constructions		Examples
sorting (subcategorizing)	property modification constructions	material term	铁 门 tiŋ <sup>2-5</sup> (muŋ <sup>55</sup> ) 'iron (door)'
		color term	乌 马 ɔu <sup>33</sup> (be <sup>52</sup> ) 'black (horse)'
		shape term	圆形 ɿ <sup>55-11</sup> -heŋ <sup>55</sup> 'round-shape, circle'
		age term	老 车 lau <sup>24-11</sup> (tɕ <sup>h</sup> ia <sup>33</sup> ) 'old (car)'
		value term	好 人 ho <sup>52-11</sup> (naŋ <sup>55</sup> ) 'good (person)'
		dimension term	薄 簿 pɔŋ <sup>55-2</sup> (p <sup>h</sup> ɔu <sup>24</sup> ) 'thin (book)'
		physical property term	硬 壳 ŋɛ <sup>24-11</sup> (k <sup>h</sup> ak <sup>2</sup> ) 'hard (shell)'

Table 6.1: CT Property Modification Constructions, adapted from Croft (forthcoming, p. 95)

Per Croft, the most prototypical property modifiers are those of the subtypes age, value, and dimension, and the next most prototypical ones are those of color and shape subtypes (Croft, forthcoming, p. 96). The ordering of the subtypes in Table 6.1 is relevant in some languages as it represents a scale where the property concepts at the top are more likely to recruit object modifier constructions while the property concepts at the bottom are more likely

<sup>57</sup> Croft (forthcoming) calls these “adjectives”.

to recruit action modifier constructions (Croft, forthcoming, p. 96, citing Wetzler, 1992).

However, these distinctions may not be relevant in Cambodian Teochew, as object modification and action modification are performed by the same strategies (see §6.3 and §6.4).

Overall, the number of tokens of CT Property Modification Constructions in my gathered data was quite low. Strategies for encoding CT Property Modification Constructions primarily include juxtaposition and compounding. Like many linguistic terms, the word “compound” or compounding is often used differently by different scholars. Per Croft (forthcoming, p. 103) the difference between juxtaposition and compounding is whether the two (or more) items are morphologically free or acting as a single word, respectively. Properties undergo tone sandhi in both strategies.

In the juxtaposition strategy, the property and the objR occur next to each other. The examples in (1) show CT Property Modification Constructions from the current data set using the juxtaposition strategy.

- (1) a. 大 番梨  
tua<sup>11</sup> huan<sup>33</sup>lai<sup>55</sup>  
big pineapple  
'big pineapple' [G1F1-9]
- b. 好 家庭  
ho<sup>52-11</sup> ke<sup>33</sup>teŋ<sup>55</sup>  
good family  
'happy family' [G3F2-4]
- c. 青 团  
te<sup>h33</sup> kiã<sup>52</sup>  
young child  
'young child' [G1F1-11]

- |    |                      |                                 |
|----|----------------------|---------------------------------|
| d. | 薄                    | 簿                               |
|    | pɔʔ <sup>5-2</sup>   | p <sup>h</sup> ɔu <sup>24</sup> |
|    | thin                 | book                            |
|    | ‘thin book’          | [G1F4-PS]                       |
| e. | 厚                    | 簿                               |
|    | kau <sup>24-11</sup> | p <sup>h</sup> ɔu <sup>24</sup> |
|    | thick                | book                            |
|    | ‘thick book’         | [G1F4-PS]                       |

In the compounding strategy, the property concept and the objR are put together to form one morphologically bound word. Whether a concept constitutes a single word is language-specific. Some characteristics of compounds are that they might have a different meaning without one of the elements, or they might be a generally accepted term for that referent in that particular language. This concept is also relevant for typifying object modification constructions explored in §6.3. The analysis of compounds using in this dissertation is based on intuitions from speaker consultants, as well as dictionaries such as Goh (2020) and TCKnow LLC. (2015).

The phrases in (2) showcase the juxtaposition strategy in CT Property Modification Constructions with the property 老 *lau*<sup>24</sup> ‘old’, while the examples in (3) showcase the compounding strategy with the same property. While there exist ‘cars’ and ‘shoes’ that are not ‘old’, ‘monkey’ and ‘tiger’ would not commonly be referred to without the modifying 老 *lau*<sup>24</sup> ‘old’. The two morphemes are bound together. However, in those words the 老 *lau*<sup>24</sup> ‘old’ has lost its semantic meaning. Instead, the words ‘monkey’ and ‘tiger’ are made up of two concepts that have merged to create one single concept. The examples in (4), which were all elicited from G2F1US and G2M2US, illustrate how 老猴 *lau*<sup>24-11</sup>-*kau*<sup>55</sup> ‘monkey’ represents a

singular concept and cannot be split by any other modifiers or intensification, and can itself be modified with the property the 老 *lau*<sup>24</sup> ‘old’.

- (2) a. 老 车  
lau<sup>24-11</sup> tɕ<sup>h</sup>ia<sup>33</sup>  
old car  
‘old car’ [G1F3-PS, G3F2-PS]
- b. 老 鞋  
lau<sup>24-11</sup> ɕi<sup>55</sup>  
old shoe  
‘old shoe(s)’ [G1F3-PS, G3F3-PS]
- (3) a. 老猴  
lau<sup>24-11</sup>-kau<sup>55</sup>  
old-monkey  
‘monkey’ [G1F1-11, G1M1-7, G1F4-PS, G3F2-PS, G3F3-PS]
- b. 老虎  
lau<sup>24-11</sup>-hou<sup>52</sup>  
old-tiger  
‘tiger’ [G1F1-10, G1M1-7, G2F1US-18]

- |     |    |  |    |   |
|-----|----|--|----|---|
| (4) | a. | 白 鞋<br>peŋ <sup>5-2</sup> ɔi <sup>55</sup><br>white shoe<br>'white shoe(s)'  | b. | 白 老猴<br>peŋ <sup>5-2</sup> lau <sup>24-11</sup> -kau <sup>55</sup><br>white old-monkey<br>'white monkey'  |
|     | c. | 老 鞋<br>lau <sup>24-11</sup> ɔi <sup>55</sup><br>old shoe<br>'old shoe(s)'  | d. | 老 老猴<br>lau <sup>24-11</sup> lau <sup>24-11</sup> -kau <sup>55</sup><br>old old-monkey<br>'old monkey'  |
|     | e. | 老 白 鞋<br>lau <sup>24-11</sup> peŋ <sup>5-2</sup> ɔi <sup>55</sup><br>old white shoe<br>'old white shoe(s)'                       | f. | 老 白 老猴<br>lau <sup>24-11</sup> peŋ <sup>5-2</sup> lau <sup>24-11</sup> -kau <sup>55</sup><br>old white old-monkey<br>'old white monkey'                       |
|     | g. | 老 老 个 鞋<br>lau <sup>24-11</sup> lau <sup>24</sup> kai <sup>55-11</sup> ɔi <sup>55</sup><br>old RED LW shoe<br>'very old shoe(s)' | h. | 老 老 个 老猴<br>lau <sup>24-11</sup> lau <sup>24</sup> kai <sup>55-11</sup> lau <sup>24-11</sup> -kau <sup>55</sup><br>old RED LW old-monkey<br>'very old monkey' |

Some other CT Property Modification Constructions using the compounding strategy are given in (5) where again two or more concepts have compounded to create a singular morphologically bound concept. These are not limited to two morphemes/concepts, as evidenced by the example in (5e) that shows a multi-property compound.

- |     |    |   |
|-----|----|---|
| (5) | a. | 长车<br>tɕ <sup>h</sup> iaŋ <sup>55-11</sup> -tɕ <sup>h</sup> ia <sup>33</sup><br>long-car<br>'bus' [G1F1-12] |
|     | b. | 青蛙<br>tɕ <sup>h</sup> ẽ <sup>33</sup> -ua <sup>33</sup><br>green-frog<br>'frog' [G1M1-7, G1F1-11]           |

- c. 凝衫  
 ŋaŋ<sup>55-11</sup>-sã<sup>33</sup>  
 cold-shirt  
 ‘coat’ [G3F1-2]
- d. 酸柑  
 suŋ<sup>33</sup>-kã<sup>33</sup>  
 sour-orange  
 ‘lime’ [G1F5US]
- e. 青红灯  
 tɕ<sup>h</sup>ẽ<sup>33</sup>-aŋ<sup>55-11</sup>-teŋ<sup>33</sup>  
 green-red-light  
 ‘traffic light’ [G1F1-12]

In some instances, when a CT Property modifies a CT ObjR, the resulting compound is another CT Property, as in (6).

- (6) a. 好食  
 ho<sup>52-24</sup>-tɕiaŋ<sup>5</sup>  
 good-food  
 ‘tasty’ [G2M1-5]
- b. 平样  
 pẽ<sup>55-11</sup>-iõ<sup>11</sup>  
 even-shape  
 ‘same’ [G1F1, G1F3, G3F1, G3F2, G3F3, G2F1US]

So far all of the given CT Property Modification Constructions have shown the order Property-ObjR, where the modifying property precedes the referent. There is at least one example of a productive suffix in Cambodian Teochew, 𠵼 *-kiã<sup>52</sup>* ‘small’, which results in an ObjR-Property order. When 𠵼 *kiã<sup>52</sup>* occurs alone it has the meaning ‘child’, as shown in (7).

- (7) a. Ø 行 觅 伊 个 囡  
 Ø kia<sup>55-11</sup> tɕ<sup>h</sup>ue<sup>11</sup> i<sup>33</sup> kai<sup>55-11</sup> kiã<sup>52</sup>  
 3SG leave look for 3SG LW child  
 ‘[She] leaves to look for her child.’ [G3F1-2]
- b. 伊 囡 昂 啊  
 i<sup>33</sup> kiã<sup>52</sup> ŋaŋ<sup>33</sup> a<sup>55</sup>  
 IK child look.up PRT  
 ‘The child looks up.’ [G1F1-11]
- c. 囡 伊 笑  
kiã<sup>52</sup> i<sup>33</sup> tɕ<sup>h</sup>io<sup>21</sup>  
child 3SG smile  
 ‘[The] child, he smiles.’ [G1F2-13]

Because of the meaning shown in (7), 囡 *kiã<sup>52</sup>* ‘child’ could be analyzed as the head of the constructions shown in (8) below, with an ObjR acting as the modifier. Here, however, we will analyze 囡 *-kiã<sup>52</sup>* ‘small’ as a modifier that is a very productive compounding strategy used to indicate small-ness of the ObjR, such as in the examples in (8).

- (8) a. 刀囡  
 tɔ<sup>33</sup>-kiã<sup>52</sup>  
 knife-small  
 ‘small knife’ [G1F3, G1F4-PS]
- b. 狗囡  
 kau<sup>52-24</sup>-kiã<sup>52</sup>  
 dog-small  
 ‘small dog; puppy’ [G1M1-7]
- c. 鸟囡  
 tɕiau<sup>52-11</sup>-kiã<sup>52</sup>  
 bird-small  
 ‘baby bird’ [G3F1-3]

- d. 丈夫囡  
 ta<sup>33</sup>pəu<sup>33</sup>-kiã<sup>52</sup>  
 man-small  
 ‘boy’ [G1F3-15, G2F1US-PS]
- e. 媪妯囡  
 tɛa<sup>33</sup>bəu<sup>52-24</sup>-kiã<sup>52</sup>  
 woman-small  
 ‘girl’ [G1F1-10, G1M1-7, G3F1-3]
- f. 草囡  
 tɛ<sup>h</sup>au<sup>52-24</sup>-kiã<sup>52</sup>  
 grass-small  
 ‘herbs’ [G1F5US-23]

The example in (9) shows the parallel meanings of 囡 -kiã<sup>52</sup> ‘small’ and 细 sɔi<sup>21</sup> ‘small.’

- (9) ... 在 许 山囡.. 细山  
 ... tɔ<sup>11</sup> hu<sup>52</sup> suã<sup>33</sup>kiã<sup>52</sup>.. sɔi<sup>21-53</sup>suã<sup>33</sup>  
 ... LOC that mountain-small.. small mountain  
 ‘...on that small mountain..small mountain’ [G1F1-11]

Finally, CT Property Modification Constructions can also use the linker strategy, where the relationship between the modifier and the objR is expressed with a certain morpheme, a linker (Croft, forthcoming, p. 123). In Cambodian Teochew, we will call 个 *kar<sup>55</sup>* a linker word, even though it is traditionally considered a classifier. As we will show in §6.2.1, we distinguish between the classifier strategy, which is a type of indexing strategy where a feature of the head objR is part of the classifier, and the linker strategy, where no such indexing is found. Furthermore, we will show that 个 *kar<sup>55</sup>* is found in many more types of constructions than any of the other classifiers. For example, 个 *kar<sup>55</sup>* can be used in Property Modification, Object Modification, and Action Modification Constructions.

There was only one example of a CT Property Modification Construction using the linker strategy in the current data set, shown in (10).

- (10) 只 个 新 个 蜀代 人  
 tɕi<sup>52</sup> kai<sup>55-11</sup> siŋ<sup>33</sup> kai<sup>55-11</sup> tɕek<sup>5-2</sup>tɔ<sup>11</sup> naŋ<sup>55</sup>  
 this LW new LW generation people  
 ‘...this new generation of people.’ [G1F7-24]

There is nothing special about the Property 新 *siŋ*<sup>33</sup> ‘new’ that would require it to be used with a linker strategy. There are examples of CT Property Modification Constructions with 新 *siŋ*<sup>33</sup> ‘new’ using the juxtaposition strategy in (11) and the compounding strategy in (12).

- (11) a. 新 鞋  
 siŋ<sup>33</sup> ɔi<sup>55</sup>  
 new shoe  
 ‘new shoe(s)’ [G3F2-PS, G3F3-PS]
- b. 新 车  
 siŋ<sup>33</sup> tɕ<sup>h</sup>ia<sup>33</sup>  
 new car  
 ‘new car’ [G3F2-PS]

- (12) 新闻  
 siŋ<sup>33</sup>-buŋ<sup>24</sup>  
 new-news  
 ‘news’ [G1F1-8]

Though there was only one example of a CT Property Modification Construction using the linker strategy, it is likely that is a broadly accepted strategy for such constructions in Cambodian Teochew. The choice of using the juxtaposition or linker strategy might come down

to speaker preference. However, more data and discussions with speakers are needed to confirm this hypothesis. According to Low (2014, pp. 26-7) and Yeo (2011, p. 30), the linker strategy is commonly used in Singapore Teochew, with examples shown in (13).

- (13) a.    tsek<sup>2</sup>    kɔ̃<sup>33</sup>                    sɔi<sup>53</sup>    sɔi<sup>213</sup>    kai<sup>11</sup>    iɔŋ<sup>2</sup>                    paŋ<sup>55</sup>  
           one    CLF.BUILDING    small    RED    LW    medicine    room  
           ‘a small clinic’ (Low, 2014, p. 27)
- b.    tseik<sup>2</sup>    tsəŋ<sup>33</sup>                    tʷa<sup>11</sup>    kai<sup>11</sup>    ts<sup>h</sup>ɪu<sup>11</sup>  
           one    CLF.TREE            be.big    LW    tree  
           ‘One big tree.’ (Yeo, 2011, p. 31)

When properties such as those in Table 6.1 modify referents, we have property modification constructions, which have already been explored in this section. However, these same properties can also occur in property reference constructions (see §5.4) or property predication constructions (see §7.3). Some general things can be said about the semantics of properties, regardless of their function, as covered in Croft (forthcoming, pp. 96-8).

The property concepts in Table 6.1 are generally understood to be stable. They represent a one-dimensional feature of a referent and thus can be said to be scalar concepts, expressing one point on a scale. There are two types of scales: antonyms, a gradient scale (‘heavy’, ‘light’); and complementaries, a “scale” with no gradience, only categorical values (‘dead’, ‘alive’) (Croft, forthcoming, p. 97). Antonyms allow for scalar modifiers (also called admodifier constructions by Croft) such as intensifiers (‘very long’), downtoners (‘somewhat long’), and superlatives (‘most expensive’), while complementaries typically do not (Croft, forthcoming, pp. 97-8). As with all categories, the use of scalar modifiers will be language-specific. There were some examples of properties being modified with scalar modifiers when

they were performing the predicate function, such as in (14). These will be covered in more detail in §7.3.

- (14) 花 数 绝. Ø 儒 绝  
hue<sup>33</sup> tɕɔi<sup>11</sup> tɕɔŋ<sup>5</sup>. Ø dzu<sup>55-11</sup> tɕɔŋ<sup>5</sup>  
flower many very. 3PL exquisite very  
‘There are a lot of flowers. [They are] very exquisite.’ [G1F4-16]

Given the relatively few tokens of properties in all functions in the current data set, more research should be done in this area to find more admodifier constructions and determine how antonyms and complementaries might be treated differently in Cambodian Teochew. For example, Xu (2007, p. 220-1) noted some differences in Jieyang in what she calls “gradable” and “ungradable” property concepts, which seem to correspond to Croft’s antonyms and complementaries, respectively. While gradable properties have a lexical counterpart/antonym, negative meanings of ungradable properties are morphologically expressed such as in the complementaries 公平 *koŋ<sup>33</sup>p<sup>h</sup>eŋ<sup>55</sup>* ‘fair’ and 唔公平 *m<sup>35-21</sup>-koŋ<sup>33</sup>p<sup>h</sup>eŋ<sup>55</sup>* ‘not fair’ (Xu, 2007, p. 220). There is also a difference in negation in Jieyang Property Predication Constructions where 𠵿 *boi<sup>35</sup>* ‘NEG-can’ is used for gradable property concepts, while 唔 *m<sup>35</sup>* ‘NEG’ is used for ungradable property concepts (Xu, 2007, p. 221-2). Some examples of negation of CT Property Predication Constructions are covered in §7.3 but more data is needed to come to more concrete conclusions.

## 6.2 Selecting modifiers

The modification subfunction of selecting selects particular instance(s) or a set of instances of a referent. Selecting is more temporary than a prototypical property (Croft, forthcoming, p. 100) (compare 三本簿 *sã<sup>33</sup>puiŋ<sup>52-24</sup>p<sup>h</sup>ɔu<sup>24</sup>* ‘three books’ to 厚簿 *kau<sup>24-11</sup>p<sup>h</sup>ɔu<sup>24</sup>*

‘a thick book’). Table 6.2 showcases the types of constructions that fulfill this selecting subfunction including enumeration modification constructions and mensural terms (§5.2.1) and quantifiers (§5.2.2). Croft (forthcoming) listed other types of selecting constructions such as set member terms with ordinal number like ‘the second box’. There were no ordinal numbers in the current data set so it represents an area for future research.

Information packaging function	Constructions		Examples
selecting	enumeration modification construction	classifier strategy	三本簿 sā <sup>33</sup> puŋ <sup>52-24</sup> p <sup>h</sup> ɔu <sup>24</sup> ‘three books’
		linker strategy	三个簿 sā <sup>33</sup> kai <sup>55-11</sup> p <sup>h</sup> ɔu <sup>24</sup> ‘three books’
		juxtaposition strategy	蜀猫 tɕek <sup>5-2</sup> ŋiau <sup>33</sup> ‘one cat’
	mensural	measure term	打(卵) ta <sup>52-24</sup> (nuŋ <sup>24</sup> ) ‘a dozen (eggs)’
		container term	蜀盒(卵) tɕek <sup>5-2</sup> ap <sup>5-2</sup> (nuŋ <sup>24</sup> ) ‘a carton (of eggs)’
		group term	撮人 tɕe <sup>h</sup> ɔɿ <sup>2-5</sup> naŋ <sup>55</sup> ‘those people’ 块人 kɔ <sup>21-53</sup> naŋ <sup>55</sup> ‘those people’
	quantifier	vague numeral	人敝 naŋ <sup>55</sup> tɕɔi <sup>11</sup> ‘many people’
		amount term	敝谢 tɕɔi <sup>11</sup> -sia <sup>11</sup> ‘many thanks’
		distributive quantifier	每年 mue <sup>52-24</sup> ni <sup>55</sup> ‘every year’

Table 6.2: CT Constructions of the selecting subfunction of the modification function, adapted from Croft (2021, p. 95)

### 6.2.1 CT Enumeration Modification Constructions

CT Cardinal Enumeration Constructions (ECs) are given in Table 6.3. The meanings ‘one’ and ‘two’ each have two different forms whose use depends on the construction.  $一 ik^2$  and  $二 dzr^{24}$  are used for the ones digits in the CT EC as in  $十一 tcap^{5-2} ik^2$  ‘eleven’ and  $蜀百一 tcek^{5-2} pe\tilde{r}^{2-5} ik^2$  ‘one hundred and one’, and  $十二 tcap^{5-2} dzr^{24}$  ‘twelve’.  $蜀 tcek^5$  and  $二 n\sigma^{24}$  are used in the beginning of the large CT Enumeration Constructions as in  $蜀百 tcek^{5-2} pe\tilde{r}^2$  and  $蜀千 tcek^{5-2} t\check{c}^h a\tilde{r}^{33}$ , and  $二千 n\sigma^{24-11} t\check{c}^h a\tilde{r}^{33}$ . Additionally, as will be shown below,  $蜀 tcek^5$  and  $二 n\sigma^{24}$  are used as modifiers in CT Enumeration Modification Constructions such as  $二人 n\sigma^{24-11} nar^{55}$  ‘two people’.<sup>58</sup>

---

<sup>58</sup> According to TCKnow LLC. (2015),  $二 dzr^{24}$  and a third form of ‘one’  $么 iau^{33}$  are used for stating phone numbers but this usage is unconfirmed for Cambodian Teochew.

Form	Meaning
空 k <sup>h</sup> aŋ <sup>21</sup>	zero
蜀 tɕek <sup>5</sup> — ik <sup>2</sup>	one
二 nɔ <sup>24</sup> 二 dʒi <sup>24</sup>	two
三 sã <sup>33</sup>	three
四 si <sup>21</sup>	four
五 ŋɔu <sup>24</sup>	five
六 lak <sup>5</sup>	six
七 tɕ <sup>h</sup> ik <sup>2</sup>	seven
八 pɔi <sup>ʔ2</sup>	eight
九 kau <sup>52</sup>	nine
十 tɕap <sup>5</sup>	ten
十一 tɕap <sup>5-2</sup> ik <sup>2</sup>	eleven
十二 tɕap <sup>5-2</sup> dʒi <sup>24</sup>	twelve
二十 dʒi <sup>24-11</sup> tɕap <sup>5</sup>	twenty
蜀百 tɕek <sup>5-2</sup> pe <sup>ʔ2</sup>	one hundred
蜀百一 tɕek <sup>5-2</sup> pe <sup>ʔ2-5</sup> ik <sup>2</sup>	one hundred one
蜀千 tɕek <sup>5-2</sup> tɕ <sup>h</sup> ai <sup>33</sup>	one thousand
两千 nɔ <sup>24-11</sup> tɕ <sup>h</sup> ai <sup>33</sup>	two thousand
蜀万 tɕek <sup>5-2</sup> buaŋ <sup>11</sup>	ten thousand

Table 6.3: Some CT Cardinal Enumeration Constructions<sup>59</sup>

Enumeration constructions can be modified by a variety of admodifiers that increase, decrease, or approximate cardinality (Croft, forthcoming, p. 98). No such constructions were found in the current data set on Cambodian Teochew so it represents an area of future research.

CT Enumeration Modification Constructions, where the EC modifies (selects) an ObjR, typically use a classifier strategy with enumeration classifiers. Enumeration classifiers identify

---

<sup>59</sup> Most of the CT Enumeration Constructions in Table 6.3 were elicited from G3F2. Variation may exist for different speakers in terms of the formation of complex numerals. More research should be done in this area.

the same referent as the objR, thus the classifier strategy is a type of indexing strategy (Croft, forthcoming, pp. 114-5). Croft (forthcoming) illustrates this relationship in Figure 6.1.

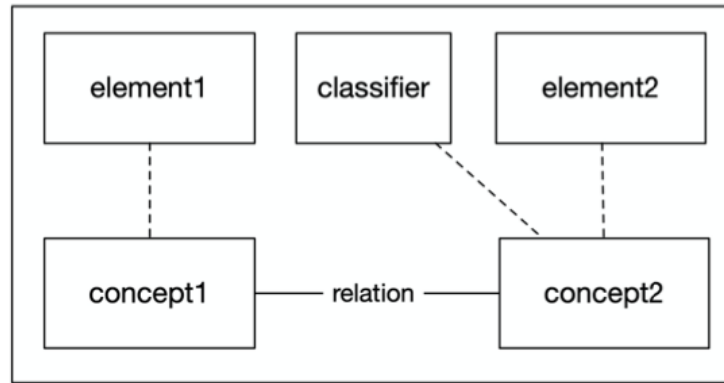


Figure 6.1: The classifier indexical strategy (Croft, forthcoming, p. 115)

In Cambodian Teochew, a variety of different classifiers exist and they are used with certain ObjRs. Examples of CT Classifiers found in the present data set along with their usage and accompanying CT ObjRs are provided in Table 6.4.

Classifier	Usage	Example ObjRs
尾 <i>bue</i> <sup>52</sup>	fish	鱼 <i>hu</i> <sup>55</sup> ‘fish’
粒 <i>liap</i> <sup>5</sup>	round things	球 <i>kiu</i> <sup>55</sup> ‘ball’ 苳𣎵 <i>lɔ<sup>11</sup>hɔŋ</i> <sup>55</sup> ‘papaya’ 𣎵𣎵 <i>maŋ<sup>11</sup>k<sup>h</sup>u<sup>ɔ</sup></i> <sup>2</sup> ‘mangosteen’ 𣎵𣎵 <i>sau</i> <sup>33</sup> <i>mau</i> <sup>33</sup> ‘rambutan’
间 <i>kaɪ</i> <sup>33</sup>	buildings	厝 <i>tɕ<sup>h</sup>u</i> <sup>21</sup> ‘house’
本 <i>puŋ</i> <sup>52</sup>	books	簿 <i>p<sup>h</sup>ɔu</i> <sup>24</sup> ‘book’ 书 <i>tɕu</i> <sup>33</sup> ‘book’
双 <i>saŋ</i> <sup>33</sup>	pair of shoes	鞋 <i>ɔɪ</i> <sup>55</sup> ‘shoe’
丛 <i>tɕaŋ</i> <sup>55</sup>	trees	树 <i>tɕ<sup>h</sup>iu</i> <sup>11</sup> ‘tree’
隻 <i>tɕia</i> <sup>ɔ</sup>	large objects, animals	老猴 <i>lau</i> <sup>24-11</sup> <i>kau</i> <sup>33</sup> ‘monkey’ 狗 <i>kau</i> <sup>52</sup> ‘dog’ 鸡 <i>kɔɪ</i> <sup>33</sup> ‘chicken’ 猪 <i>tu</i> <sup>33</sup> ‘pig’ 马 <i>be</i> <sup>52</sup> ‘horse’ 鹿 <i>tek</i> <sup>5</sup> ‘deer’
条 <i>tiau</i> <sup>55</sup>	long items	桥 <i>kiɔ</i> <sup>55</sup> ‘bridge’
张 <i>tiɔ</i> <sup>33</sup>	flat things	图 <i>tɔu</i> <sup>55</sup> ‘picture’
对 <i>tui</i> <sup>21</sup>	pair of humans	翁妪 <i>aŋ</i> <sup>33</sup> <i>bɔu</i> <sup>52</sup> ‘couple’

Table 6.4: Sortal classifiers in Cambodian Teochew

Canonical word order for CT Enumeration Modification Constructions using the classifier strategy is EC + Classifier + ObjR, as in (15).

- (15) a. 三 尾 鱼  
*sã*<sup>33</sup> *bue*<sup>52-11</sup> *hu*<sup>55</sup>  
 three CLF.FISH fish  
 ‘three fish’ [G3F1, G3F2, G3F3-PS]
- b. 二 间 厝  
*nɔ*<sup>24-11</sup> *kaɪ*<sup>33</sup> *tɕ<sup>h</sup>u*<sup>21</sup>  
 two CLF.BUILDING house  
 ‘two houses’ [G1F4, G2F2-PS]

- c. 三 本 簿  
sã<sup>33</sup> puŋ<sup>52-24</sup> p<sup>h</sup>ɔu<sup>24</sup>  
three CLF.BOOK book  
‘three books’ [G1F1, G1F2, G1F3, G1F4-PS]
- d. 二 双 鞋  
nɔ<sup>24-21</sup> saŋ<sup>33</sup> ɔi<sup>55</sup>  
two CLF.PAIR.SHOE shoe  
‘two pairs of shoes’ [G3F3-PS]
- e. 蜀 隻 狗  
tɕek<sup>5-2</sup> tɕiaɪ<sup>2-5</sup> kau<sup>52</sup>  
one CLF.ANIMAL dog  
‘one dog’ [G1F4-PS]
- f. 三 对 翁妪  
sã<sup>33</sup> tui<sup>21-53</sup> aŋ<sup>33</sup>bɔu<sup>52</sup>  
three CLF.PAIR.HUMAN husband-wife  
‘three couples’ [G1F4, G3F2-PS]

The classifier strategy can also be used with loanwords (see Chapter 8), at least for some speakers. As Table 6.4 shows, words borrowed from Khmer that meet the usage requirements are allowed with the classifier. For example, the classifier 粒 *liap*<sup>5</sup> used for round-ish items can be used for historic Teochew words like 球 *kiu*<sup>55</sup> ‘ball’ in (16a), as well as with 莽吉朥 *maj*<sup>11</sup>*k<sup>h</sup>u*<sup>2</sup> ‘mangosteen’ in (16b), and 荔枝 *sau*<sup>33</sup>*mau*<sup>33</sup> ‘rambutan’ in (16c).

- (16) a. 六 粒 球  
lak<sup>5-2</sup> liap<sup>5-2</sup> kiu<sup>55</sup>  
six CLF.ROUND ball  
‘six balls’ [G1F4, G3F3-PS]

- b. 三 粒 មង្គុត  
sã<sup>33</sup> liap<sup>5-2</sup> maŋ<sup>11k<sup>h</sup>u<sup>?</sup>2</sup>  
three CLF.ROUND mangosteen  
‘three mangosteens’ [G1F4-PS]
- c. 二 粒 សាវម៉ារី  
nɔ<sup>25-11</sup> liap<sup>5-2</sup> sau<sup>33</sup>mau<sup>33</sup>  
two CLF.ROUND rambutan  
‘two rambutans’ [G1F3-PS]

The type of classifiers presented in Table 6.4 are examples of sortal classifiers, where some feature of the ObjR is coded/indexed in the classifier (Croft, forthcoming, pp. 123-3). Cambodian Teochew also occasionally recruits the classifier strategy for CT Mensural Constructions through the use of mensural classifiers. These constructions differ from the sortal classifiers because the classifier itself does not relate to any inherent property of the ObjR. Few examples of such classifiers were found in the data set, but they include the measure term 打 *ta*<sup>52</sup> ‘dozen’, in (17), the container term 盒 *ap*<sup>5</sup> ‘box’ in (18), and the group terms, 撮 *tɕ<sup>h</sup>ɔ<sup>?</sup>2* and 块 *kɔ<sup>?</sup>1*, which also indicate plurality of a referent, as in (19). Future research should further investigate the use and existence of these types of mensural terms.

- (17) 打 卵  
ta<sup>52-24</sup> nuŋ<sup>24</sup>  
dozen egg  
‘a dozen eggs’ [G1F1-PS]

- (18) 蜀 盒 卵  
tɕek<sup>5-2</sup> ap<sup>5-2</sup> nuŋ<sup>24</sup>  
one box egg  
‘a carton of eggs’ [G3F3-PS]

- (19) a. 伊 睇见 只 撮 鸟团 □ 鸟宿  
 i<sup>33</sup> t<sup>h</sup>ɔ̃<sup>52-11</sup> kian<sup>21-53</sup> tɕi<sup>52</sup> tɕ<sup>h</sup>ɔ̃<sup>2-5</sup> tɕiau<sup>52-11</sup>-kiã<sup>52</sup> na<sup>11</sup> tɕiau<sup>52-11</sup>-siu<sup>21</sup>  
 3SG see this CLF.GROUP bird-baby LOC bird-nest  
 ‘He sees these baby birds in a bird nest.’ [G3F1-3]
- b. 有 块 人 在.. 在 艇\* 船  
 u<sup>11</sup> kɔ̃<sup>21-53</sup> nan<sup>55</sup> tɔ̃<sup>11</sup>.. tɔ̃<sup>11</sup> p<sup>h</sup>e<sup>24-11</sup> tɕuŋ<sup>55</sup>  
 have CLF.GROUP person PROG.. PROG paddle boat  
 ‘There are people paddling a boat.’ [G1F4-16]

Xu (2007) listed 双 *saj*<sup>33</sup> ‘pair’ and 对 *tui*<sup>33</sup> ‘couple’ as “collective classifiers”, a type of mensural classifier (p. 99). Though these terms are slightly different from the other sortal classifiers in Table 6.4 in that they denote the meaning of a pair or couple, in Cambodian Teochew they appear to have limited use and still code a property of the ObjR. Thus, although the two terms both denote the meaning of a pair in (15d) and (15f), it is unacceptable to interchange them, as shown by the ungrammaticality of the examples in (20). Additionally, as will be shown later in (43), the head ObjR can be elided in CT Enumeration Modification Constructions with 双 *saj*<sup>33</sup> and 对 *tui*<sup>21</sup> because some characteristics of the ObjR are indexed on the classifier. For these reasons I have classified 双 *saj*<sup>33</sup> and 对 *tui*<sup>21</sup> as sortal classifiers in Cambodian Teochew.

- (20) a. 二 对 鞋  
 \*nɔ̃<sup>24-21</sup> tui<sup>21-53</sup> ɔ̃i<sup>55</sup>  
 two CLF.PAIR shoe  
 ‘\*two pairs of shoes’
- b. 三 双 翁妘  
 \*sã<sup>33</sup> saŋ<sup>33</sup> aŋ<sup>33</sup>bɔu<sup>52</sup>  
 three CLF.PAIR husband-wife  
 ‘\*three couples’

Per Croft (forthcoming, p. 116), the classifier strategy can extend beyond enumeration modification constructions according to the hierarchy of modifier types in (21).

(21) ECs, quantifiers < deixis < property < action, object

There is some evidence that CT Quantifier Constructions (covered in more detail in §6.2.2) use the classifier strategy, as in (22).

(22) 只个 拢是 丛 树  
 tɕei<sup>55</sup> loŋ<sup>52-24</sup>si<sup>11</sup> tɕaŋ<sup>11</sup> tɕ<sup>h</sup>iu<sup>11</sup>  
 PDCO all CLF.TREE tree  
 ‘This (one [the picture]) [is] all trees.’ [G1F1-10]

CT Deictic Contextual Attribute Constructions<sup>60</sup> can also use the classifier strategy, shown in (23).

(23) a. 只 条 路 ...  
 tɕei<sup>52</sup> tiau<sup>55-11</sup> lou<sup>11</sup> ...  
 this CLF.LONG road ...  
 ‘This road...’ [G2F1US-19]

b. □ 只 张 图 底 □ ...  
 na<sup>11</sup> tɕei<sup>52</sup> tiʃ<sup>33</sup> tou<sup>55-11</sup> toi<sup>52</sup> le<sup>33</sup> ...  
 at this CLF.FLAT picture inside PRT ...  
 ‘Inside this picture...’ [G1F4-16]

c. 只 丛 石榴 ...  
 tɕei<sup>52</sup> tɕaŋ<sup>11</sup> siɔŋ<sup>5-2</sup>liu<sup>55</sup> ...  
 this CLF.TREE pomegranate ...  
 ‘This pomegranate tree...’ [G1F5US-23]

---

<sup>60</sup> Note that CT Deictic Contextual Attribute Constructions perform the modification subfunction of situating. They are covered here as their use of the classifier, linker, and juxtaposition strategies parallel those used in CT Enumeration Modification Constructions.

All of the example so far have shown the canonical word order of EC/Quantifier/Deixis + Classifier + ObjR. However, there were several cases where word order surfaced as ObjR + EC + Classifier for speakers G1F3 and G1F4, as in (24). According to Chao (1968/2011, pp. 570-1) this order is accepted in Chinese in contexts where one is taking inventory of items. This does not seem to be the same context that we have here. The examples in (24) came from consultants describing pictures where there was only one type of item. So for (24a), the picture was of four identical dogs, with no other items to take “inventory” of in the picture.

- (24) a. 狗 四 隻  
 kau<sup>52</sup> si<sup>21-53</sup> tɕiaŋ<sup>2</sup>  
 dog four CLF.ANIMAL  
 ‘four dogs’ [G1F3-PS]
- b. ល្អិត មួយ គ្រាប់  
 lo<sup>11</sup> hɔŋ<sup>55</sup> tɕek<sup>5-2</sup> liap<sup>5</sup>  
 papaya one CLF.ROUND  
 ‘one papaya’ [G1F4-PS]
- c. សាវម៉ែ មួយ គ្រាប់  
 sau<sup>33</sup> mau<sup>33</sup> tɕek<sup>5-2</sup> liap<sup>5</sup>  
 rambutan one CLF.ROUND  
 ‘one rambutan’ [G1F3, G1F4-PS]
- d. 只个 厝 二 间  
 tɕi<sup>52</sup> kai<sup>55</sup> tɕ<sup>h</sup>u<sup>21</sup> nɔ<sup>24-11</sup> kai<sup>33</sup>  
 PDCO house two CLF.BUILDING  
 ‘this (one [the picture]) [is] two houses’ [G1F3-PS]
- e. 鞋 二 双  
 ɔi<sup>55</sup> nɔ<sup>24-11</sup> saŋ<sup>33</sup>  
 shoe two CLF.PAIR.SHOE  
 ‘two pairs of shoes’ [G1F3-PS]

Similarly, three speakers used the non-canonical word order for the CT Mensural Construction in (25).

- (25) 卵 蜀 打  
 nuŋ<sup>24</sup> tɕek<sup>5-2</sup> ta<sup>52</sup>  
 egg one dozen  
 ‘a dozen eggs’ [G1F3-PS, G1F4-PS, G2M1-PS]

Of particular interest is the tone sandhi on words, or lack of it. In the canonical word order in (26), both the EC and the Classifier undergo tone sandhi.

- (26) 六 粒 球  
 lak<sup>5-2</sup> liap<sup>5-2</sup> kiu<sup>55</sup>  
 six CLF.ROUND ball  
 ‘six balls’ [G1F4-PS]

In the ObjR + EC + Classifier order in (27), only the EC undergoes tone sandhi. The ObjR and Classifier do not. There is no pause between the ObjR and the EC.

- (27) 球 六 粒  
 kiu<sup>55</sup> lak<sup>5-2</sup> liap<sup>55</sup>  
 ball six CLF.ROUND  
 ‘six balls’ [G1F3-PS]

There is some evidence that this order occurs with CT Deictic Contextual Attribute Constructions like the example in (28).

- (28) □ 相 底畔 有 鸡 只 隻  
 na<sup>11</sup> siɔ<sup>21</sup> tɔi<sup>52-11</sup> paɪ<sup>55</sup> u<sup>24-11</sup> kɔi<sup>33</sup> tɕi<sup>52</sup> tɕia?<sup>2</sup>  
 LOC picture inside have chicken this CLF.ANIMAL  
 ‘There is this chicken in the picture.’ [G3F1-1]

CT Enumeration Modification Constructions can also use the linker strategy, with 个 *kai*<sup>55-11</sup> ‘LW’. Almost all CT Common ObjRs can be enumerated using 个 *kai*<sup>55-11</sup> rather than any of the classifiers in Table 6.4. The examples in (29) show CT Enumeration Modification Constructions using the linker strategy.

- (29) a. 六 个 球  
*lak*<sup>5-2</sup> *kai*<sup>55-11</sup> *kiu*<sup>55</sup>  
 six LW ball  
 ‘six balls’ [G3F1, G3F2-PS]
- b. 四 个 狗  
*si*<sup>21-53</sup> *kai*<sup>55-11</sup> *kau*<sup>52</sup>  
 four LW dog  
 ‘four dogs’ [G3F2-PS]
- c. 三 个 簿  
*sã*<sup>33</sup> *kai*<sup>55-11</sup> *p<sup>h</sup>ɔu*<sup>24</sup>  
 three LW book  
 ‘three books’ [G3F1, G3F2-PS]
- d. 三 个 𦰩  
*sã*<sup>33</sup> *kai*<sup>55-11</sup> *lo<sup>11</sup>hɔŋ*<sup>55</sup>  
 three LW papaya  
 ‘three papayas’ [G3F1, G3F2-PS]
- e. 二 个 𦰩  
*nɔ*<sup>24-11</sup> *kai*<sup>55-11</sup> *maŋ<sup>11</sup>k<sup>h</sup>u?*<sup>2</sup>  
 two LW mangosteen  
 ‘two mangosteens’ [G3F1, G3F2, G3F3-PS]
- f. 三 个 翁妪  
*sã*<sup>33</sup> *kai*<sup>55-11</sup> *aŋ<sup>33</sup>bɔu*<sup>52</sup>  
 three LW husband-wife  
 ‘three couples’ [G3F1]

There were no tokens of CT Enumeration Modification Constructions using the linker strategy with the non-canonical word order. In Figure 6.2 from Croft (forthcoming) that illustrates the linker strategy for relating two concepts, the linker word appears between the two concepts it is relating. In the non-canonical word order, the linker would be in a different position and thus would not as clearly signal a relationship between the concept1 and concept2.

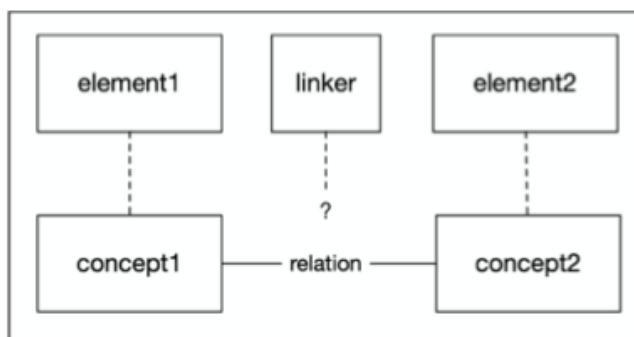


Figure 6.2: The linker strategy for relating two concepts (Croft, forthcoming, p. 123)

Some CT ObjRs can be enumerated without the Classifier or linker word. Thus, the CT Enumeration Modification Construction can also use a juxtaposition strategy. While enumerating the ObjR 人 *naŋ*<sup>55</sup> ‘people’ can use the linker strategy in (30), more often than not the juxtaposition strategy is used, as in (31).

- (30) 有 四 个 人, 四 个 人  
 u<sup>11</sup> si<sup>21-53</sup> kai<sup>55-11</sup> naŋ<sup>55</sup>, si<sup>21-53</sup> kai<sup>55-11</sup> naŋ<sup>55</sup>  
 have four LW person, four LW person  
 ‘There are four people, four people.’ [G3F2-4]
- (31) a. 伊 二 人 刻 苦  
 i<sup>33</sup> nɔ<sup>24-11</sup> naŋ<sup>55</sup> k<sup>h</sup>ak<sup>5</sup> k<sup>h</sup>ɔu<sup>52</sup>  
 IK two person work hard  
 ‘The two people work hard.’ [G3F2-4]

- b. 伊 二 人 在 倒 揽  
 i<sup>33</sup> nɔ<sup>24-11</sup> nan<sup>55</sup> tɔ<sup>11</sup> tau<sup>52-24</sup> lam<sup>52</sup>  
 IK two person PROG swap hug  
 ‘The two people are hugging.’ [G1F1-10]
- c. 只 二 人 ...  
 tɕi<sup>52</sup> nɔ<sup>24-11</sup> nan<sup>55</sup> ...  
 this two person ...  
 ‘These two people...’ [G1M1-7]

Some time-related CT ObjRs also use the juxtaposition strategy, in (32). For these words, it is generally unacceptable to use the linker strategy as in (33). 点 *tiam*<sup>52</sup> is historically a classifier itself, with the head objR 钟 *tɕeŋ*<sup>33</sup> omitted, though no evidence for this usage was found for Cambodian Teochew, while 年 *ni*<sup>55</sup> ‘year’ and 日 *dʒik*<sup>5</sup> ‘day’ are auto-classifiers (R.J. LaPolla, personal communication, December 19, 2021).

- (32) a. ... 在 只 个 二 点 就 练 到 五 点  
 ... tɔ<sup>11</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> nɔ<sup>24-11</sup> tiam<sup>52</sup> tɕiu<sup>11</sup> lian<sup>11</sup> kau<sup>21-53</sup> ŋɔu<sup>24-11</sup> tiam<sup>52</sup>  
 ... at this LW two point then practice until five point  
 ‘...[I] practiced from 2:00 to 5:00.’ [G1M2-24]
- b. ... 蜀 二 年  
 ... tɕek<sup>5-2</sup> nɔ<sup>24-11</sup> ni<sup>55</sup>  
 ... one two year  
 ‘...one to two years’ [G1F6-24]
- c. 有 蜀 日 ...  
 u<sup>11</sup> tɕek<sup>5-2</sup> dʒik<sup>5</sup> ...  
 have one day ...  
 ‘one day...’ [G1F7-24, G3F2-24, G3F2-24, G2F1-US]

- (33) a. 三 个 点  
 sã<sup>33</sup> kai<sup>55-11</sup> tiam<sup>52</sup>  
 two LW point  
 ‘\*two o’clock’
- b. 三 个 年  
 sã<sup>33</sup> kai<sup>55-11</sup> ni<sup>55</sup>  
 three LW year  
 ‘\*three years’
- c. 三 个 日  
 sã<sup>33</sup> kai<sup>55-11</sup> dzik<sup>5</sup>  
 three LW day  
 ‘\*three days’

Interestingly, not all time-related CT ObjRs exclusively use the juxtaposition strategy.

月 *gueŋ*<sup>5</sup> ‘month’ uses the linker strategy, in (34).

- (34) 三 个 月  
 sã<sup>33</sup> kai<sup>55-11</sup> gueŋ<sup>5</sup>  
 three LW month  
 ‘three months’ [G2M2-E]

So far, Cambodian Teochew, as shown in examples (30-34), seems to pattern similarly with other Teochew varieties as to whether an ObjR takes a linker or not. For example, the juxtaposition strategy of the enumeration of ObjR 人 *naŋ*<sup>55</sup> ‘people’ is also used in Singapore Teochew in (35a) and Jambi Teochew in (35b).<sup>61</sup>

---

<sup>61</sup> X. Chen (2003, p. 382) also presented the non-transcribed example below for Johor Bahru Teochew.

i. 二人扛了蜀大概柴

- (35) a. u<sup>11</sup> pian<sup>11</sup> k<sup>h</sup>ɔi<sup>ʔ2</sup> i<sup>33</sup> nɔ<sup>11</sup> nan<sup>55</sup> tsɔ<sup>53</sup> bu<sup>33</sup> t<sup>h</sup>ak<sup>2</sup> kau<sup>53</sup> liou<sup>11</sup> bɔ<sup>55</sup>  
 have way give 3SG two person together read until finish NEG.have  
 ‘Would there be a way for both of them to finish studying together?’  
 (Low, 2014, p. 46)
- b. si<sup>11</sup> ngou<sup>213</sup> nang<sup>35</sup> ...  
 four five people  
 ‘Four or five people...’ (Peng, 2012, p. 384)

The examples in (36) show how the Jieyang Enumeration Modification Constructions for *zek*<sup>5</sup> ‘day’ and *ni*<sup>55</sup> ‘year’ use the juxtaposition strategy. Xu (2007, p. 244) noted how ‘year’ and ‘day’ don’t required the linker, while *gue*<sup>ʔ5</sup> ‘month’ does.

- (36) a. ... tsek<sup>5-2</sup> zek<sup>5</sup> ...  
 ... one day ...  
 ‘...one day...’ (Xu, 2007, p. 70)
- b. zi<sup>35-21</sup> tsap<sup>5-2</sup> gou<sup>35-21</sup> ni<sup>55</sup>  
 two ten five year  
 ‘twenty-five years’ (Xu, 2007, p. 32)

Low (2014) similarly noted that the linker strategy is required for ST Enumeration Modification Constructions with *gue*<sup>ʔ5</sup> ‘month’ in (37).

- (37) tsek<sup>2</sup> kai<sup>11</sup> gue<sup>ʔ5</sup>  
 one LW month  
 ‘one month’ (Low, 2014, p. 26)

Returning to Cambodian Teochew, there is some evidence that the Classifier or linker word is not strictly required, even outside of the aforementioned words. The juxtaposition strategy was also used for enumerating some CT ObjRs where one would typically expect a classifier or linker strategy to be used such as the examples in (38).

- (38) a. 蜀 猫  
 tɕek<sup>5-2</sup> ŋiau<sup>33</sup>  
 one cat  
 ‘One/a cat.’ [G1F1-10]
- b. 蜀 猫 死 啊  
 tɕek<sup>5-2</sup> ŋiau<sup>33</sup> uk<sup>5</sup> a<sup>33</sup>  
 one cat sleep PRT  
 ‘A cat sleeps.’ [G1F2-14]
- b. 蜀 礼 鱼  
 tɕek<sup>5-2</sup> li<sup>52-24</sup> hu<sup>55</sup>  
 one special fish  
 ‘one special fish’ [G2F1US-10]
- c. 四 方向  
 si<sup>21-53</sup> huŋ<sup>33</sup> hian<sup>52</sup>  
 four directions  
 ‘four directions’ [G1F3-PS]

There were a couple instances where speakers self-corrected to include the sortal classifier in (39a) or a mensural term in (39b).

- (39) a. 二 鹰鹞 □ ... 二 隻 鹰鹞 □  
 nɔ<sup>24-11</sup> eŋ<sup>33</sup> bu<sup>52</sup> na<sup>33</sup> ... nɔ<sup>24-11</sup> tɕial<sup>ʔ2-5</sup> eŋ<sup>33</sup> bu<sup>52</sup> na<sup>33</sup>  
 two eagle PRT ... two CLF.ANIMAL eagle PRT  
 ‘Two eagles...two eagles.’ [G1M1-7]
- b. 有 蜀 柴 .. 蜀, 蜀 块 啊  
 u<sup>11</sup> tɕek<sup>5-2</sup> tɕ<sup>h</sup> a<sup>55</sup> .. tɕek<sup>5-2</sup>, tɕek<sup>5-2</sup> kɔ<sup>21</sup> a<sup>55</sup>  
 have one wood .. one, one CLF.PIECE PRT  
 ‘There is one wood..one piece [of wood.]’ [G1F1-11]

The examples in (38) show EC + ObjR word order. There are also instances of the juxtaposition strategy used with ObjR + EC word order for speaker G1F3 in (40). When the order is ObjR + EC, the numeral appears in citation form.<sup>62</sup>

- (40) a. 𠵿𠵿𠵿 二  
 maŋ<sup>11</sup>k<sup>h</sup>uŋ<sup>2</sup> nɔ<sup>24</sup>  
 mangosteen two  
 ‘two mangosteens’ [G1F3-PS]
- b. 𠵿𠵿 蜀  
 lo<sup>11</sup>hɔŋ<sup>55</sup> tɕek<sup>5</sup>  
 papaya one  
 ‘one papaya’ [G1F3-PS]
- c. 鱼 三  
 hu<sup>55</sup> sã<sup>33</sup>  
 fish three  
 ‘three fish’ [G1F3-PS]
- d. 只个 簿 三  
 tɕi<sup>52</sup>kai<sup>55</sup> p<sup>h</sup>ɔu<sup>24</sup> sã<sup>33</sup>  
 PDCO book three  
 ‘this (one [the picture]) [is] three books’ [G1F3-PS]

---

<sup>62</sup> Zhang (2013, p. 33) showed that in listing contexts in Mandarin Chinese, the classifier is not required. However, for Zhang, in order to qualify as a “list”, multiple items need to be listed, as in (i). This was not the case for Cambodian Teochew as the items in (40) were enumerated individually.

- (i) a. san fang liang ting  
 three room two sitting.room  
 ‘three bedrooms and two sitting rooms’ (Zhang, 2013, p. 33)
- b. si cai yi tang  
 four vegetable one soup  
 ‘four units of vegetable and one unit of soup’ (Zhang, 2013, p. 33)

Finally, there were also cases of CT Deictic Contextual Attribute Constructions that used the juxtaposition strategy, as in (41).

- (41) a. 只 鸟 飞  
 tɕi<sup>52</sup> tɕiau<sup>52</sup> pue<sup>33</sup>  
 this bird fly  
 ‘This bird flies.’ [G1F1-9]
- b. 伊 想 只 鸭 想 獐\*鼠  
 i<sup>33</sup> siŋ<sup>24-11</sup> tɕi<sup>52</sup> aʔ<sup>2</sup> siŋ<sup>24-11</sup> ŋiau<sup>24</sup>tɕ<sup>h</sup>u<sup>52</sup>  
 3SG think this duck think mouse  
 ‘It thinks of this duck and thinks of [a] mouse.’ [G1F1-9]
- c. 伊 欢喜 觅 挈 掠 只 獐\*鼠  
 i<sup>33</sup> huã<sup>33</sup>hi<sup>52</sup> tɕ<sup>h</sup>ue<sup>11</sup> k<sup>h</sup>iaʔ<sup>5-2</sup> liaʔ<sup>5-2</sup> tɕi<sup>52</sup> ŋiau<sup>11</sup>tɕ<sup>h</sup>u<sup>52</sup>  
 3SG happy find take catch this mouse  
 ‘He is happy to find this mouse to catch.’ [G3F1-3]

The non-canonical word orders of ObjR + EC + Classifier in (24) and ObjR + EC in (40) were not found in the literature on other Teochew varieties. These forms appear to show the influence of Khmer on Cambodian Teochew. Word order of Khmer Enumeration Modification Constructions using the classifier strategy is ObjR + EC + Classifier. However, classifiers are not obligatory, as Khmer Enumeration Modification Constructions can also use the juxtaposition strategy. This is exemplified in (42a). Khmer Mensural Constructions can similarly use either the classifier strategy or the juxtaposition strategy, as in (42b).

- (42) a. កូន ប្រាំ (នាក់ ១)  
 koon pram (neak)  
 child five CLF.PERSON  
 five children (Sak-Humphry, 2016, p. 224)
- b. សៀវភៅ បី (ក្បាល)  
 siəvp<sup>h</sup>iv bəy (kbaal)  
 book three (volume)  
 ‘3 (volumes) of books or 3 books’ (Sak-Humphry, 2016, p. 63)

The shift from EC + Classifier + ObjR order to ObjR + EC + Classifier order goes against a hypothesis in Greenberg (1990). He proposed that if both ObjR + EC + Classifier and EC + Classifier + ObjR orders<sup>63</sup> can be found in a language presently and/or historically, the existence of the former order would precede the latter (Greenberg, 1990, p. 229). He showed how Early Archaic Chinese (Dobson, 1962) exhibited ObjR-(EC-CLF) order, while Chinese of the present day has (EC-CLF)-ObjR order.<sup>64</sup> At the same time, he stated that the order in Modern Khmer of ObjR-(EC-CLF) is the same as that in Old Khmer (Greenberg, 1990, p. 272) and thus no shift has happened. For a language to first exhibit a canonical order of (EC-CLF)-ObjR and then shift to allow ObjR-(EC-CLF) would not be expected. Yet this seems to be exactly what we have here. More research should be done in this area to establish the extent of ObjR-(EC-CLF) usage amongst Cambodian Teochew speakers.

That there were several instances of CT Enumeration Modification Constructions and CT Deictic Contextual Attribute Constructions using the juxtaposition strategy likely also shows an influence from Khmer. While some ObjRs in other Teochew varieties can be

---

<sup>63</sup> Greenberg’s terminology/abbreviations have been adjusted to our terms used here.

<sup>64</sup> Greenberg even stated that “[a]ll Chinese dialects... have, as far as I know, [(EC-CLF)-ObjR]” (Greenberg, 1990, p. 229). What we are proposing here is that this Sinitic variety shows ObjR-(EC-CLF) in addition to (EC-CLF)-ObjR.

enumerated without a linker or classifier, the number of words able to surface with this form in Cambodian Teochew is higher. The self-correcting done by the speakers in (39) shows that those speakers have a conscious understanding that such constructions have traditionally required the use of the classifier strategy. However, at least for speaker G1F1, not all instances were corrected so she utters constructions without the classifier at least some of the time. I will argue that juxtaposition is a viable strategy for such CT Constructions, at least for some speakers.<sup>65</sup>

Additionally, the data I have collected is subject to the observer's paradox. As speakers were aware of my research intentions to record the Cambodian Teochew language, they may have wanted to present a more traditional way of speaking (i.e. using the classifier strategy). Future research could explore these types of constructions in more naturalistic texts such as dialogues between speakers to learn more about their occurrences.

Another thing to note about classifiers is that they can appear without the objR if it can be determined from either the linguistic or non-linguistic context (Greenberg, 1972, p. 6). This is because the classifier strategy is an indexing strategy and therefore some characteristics of the head objR are indexed on the classifier. There are examples of this behavior in Cambodian Teochew. The examples in (43) are from the picture stimuli sessions in which the head objR could be determined from the non-linguistic context (the picture). In example (44), speaker G1F1 was pointing all around the particular picture in the Zoom book she was narrating where

---

<sup>65</sup> These findings are contrary to Erbaugh's (2002) analysis of Cantonese and Mandarin tellings of The Pear Story where no storyteller (n = 49) omitted a classifier (or linker) after an enumeration or deictic word (p. 36).

there were many coconut trees. Thus the head ObjR ‘tree’ could similarly be inferred from the non-linguistic context.

- (43) a. 二 双  
 nɔ̃<sup>24-11</sup> saŋ<sup>33</sup>  
 two CLF.PAIR.SHOE  
 ‘two pairs of shoes’ [G1F4-PS]
- b. 三 对  
 sã<sup>33</sup> tui<sup>21</sup>  
 three CLF.PAIR.PEOPLE  
 ‘three couples’ [G1F3, G1F4, G3F3-PS]
- c. 蜀 双 雅, 蜀 双 歪 雅  
 tɕek<sup>5-2</sup> saŋ<sup>33</sup> ŋia<sup>52</sup>, tɕek<sup>5-2</sup> saŋ<sup>33</sup> bɔi<sup>24-11</sup> ŋia<sup>52</sup>  
 one CLF.PAIR.SHOE pretty, one CLF.PAIR.SHOE not pretty  
 ‘one pair [of shoes] is pretty, one pair [of shoes] is not pretty’ [G1F4-PS]

- (44) 只个,只个 椰子 丛 呢 ... 椰子 丛 ... 椰子 丛  
 tɕei<sup>55</sup>, tɕei<sup>55</sup> ia<sup>11</sup>tɕi<sup>52-24</sup> tɕaŋ<sup>55</sup> ne<sup>11</sup> ... ia<sup>11</sup>tɕi<sup>52-24</sup> tɕaŋ<sup>55</sup> ... ia<sup>11</sup>tɕi<sup>52-24</sup> tɕaŋ<sup>55</sup>  
 PDCO PDCO coconut CLF.TREE PRT ... coconut CLF.TREE ... coconut CLF.TREE  
 ‘this one is a coconut tree...coconut tree(s)...coconut tree(s)’ [G1F1-8]

Sometimes the same ObjR can be modified by different sortal classifiers. For example, the Classifier 丛 tɕaŋ<sup>55</sup> ‘CLF.TREE’ can be used to modify the fruit ObjR when the real world referent is that fruit’s tree, as in the first part of the utterance in (45). Then when counting the fruit of that tree, the ObjR would be modified with the classifier 粒 liap<sup>5</sup> ‘CLF.ROUND’ (assuming the fruit is round), shown in the second part of the utterance in (45).

- (45) 只个 四 丛 肉闲. 生 几 粒-团 □  
 tɕei<sup>52</sup>kai<sup>55</sup> si<sup>21-53</sup> tɕaŋ<sup>55-11</sup> nek<sup>5-2</sup>ai<sup>52</sup>. sɛ̃<sup>33</sup> kua<sup>53</sup> liap<sup>5</sup>-kiã<sup>52</sup> nia<sup>33</sup>  
 PDCO four CLF.TREE longan. fruit several CLF.ROUND-small PRT  
 ‘This (one) [is] four longan trees. There are a lot of small fruits.’ [G1F5US-23]

Overall the number of unique classifiers found in my data set was not nearly as diverse as those presented in the Jieyang variety in Xu (2007, pp. 98-106) and in Xie (2008). And there seems to be a prevalence of the use of the linker strategy rather than the classifier strategy, at least for some speakers. For example, G3F1 nearly always enumerated using the linker 个 *kai*<sup>55</sup>. G3F2 also frequently used the linker strategy. There were also two instances where she first used the linker strategy and then self-corrected to the classifier strategy upon hearing speaker G3F3 using it, as in (46a-b). G3F3 similarly did so once in (46c).

- (46) a. 三 个 簿 ... 三 本 簿  
 sā<sup>33</sup> kai<sup>55-11</sup> p<sup>h</sup>ou<sup>24</sup> ... sā<sup>33</sup> puŋ<sup>52-24</sup> p<sup>h</sup>ou<sup>24</sup>  
 three LW book ... three CLF.BOOK book  
 ‘three books...three books’ [G3F2-PS]
- b. 四 个 狗 ... 四 隻 狗  
 si<sup>21-53</sup> kai<sup>55-11</sup> kau<sup>52</sup> ... si<sup>21-53</sup> tɕiaŋ<sup>2-5</sup> kau<sup>52</sup>  
 four LW dog ... four CLF.ANIMAL dog  
 ‘four dogs...four dogs’ [G3F2-PS]
- c. 九 个- 九 隻 老猴  
 kau<sup>52-11</sup> kai<sup>55-11-</sup> kau<sup>52-11</sup> tɕiaŋ<sup>2-5</sup> lau<sup>24-11</sup> kau<sup>55</sup>  
 nine LW- nine CLF.ANIMAL monkey  
 ‘nine- nine monkeys’ [G3F3-PS]

The self-correcting done by speakers in (46), and that in (39), along with discussions with speakers, highlighted that they have a conscious understanding of the classifiers covered in this section and their usages. Nonetheless, since enumerating with the linker 个 *kai*<sup>55</sup> is generally accepted for all CT ObjRs, often times these same speakers will utter CT Enumerating Modification Constructions with the linker (or even juxtaposition) strategy. The

reduction in the number of classifiers and the increasing prevalence of the use of the linker/juxtaposition strategies, is likely an area of contact-induced language change.

A similar phenomenon was found by Schmidt (1983), who explored differences in the Dyirbal language as spoken by the older and the younger generations. Due to a decline in the number of speakers and therefore reduced input, she found that the young people speak a simplified version of Dyirbal. For example, the classifier system was cited as a target for simplification or loss in language attrition (Schmidt, 1983, p. 220). A reduction in the classifier system by younger speakers was also found for Nivkh (Gruzdeva, 2004). This follows an assertion from Romaine (2010, p. 326) that unless the input for marked forms (like indexing classifiers) reaches a certain threshold, speakers are not likely to acquire them. In the case of Cambodian Teochew, that the dominant language in the contact environment, Khmer, has fewer classifiers, and that they are not obligatory, likely contributes to this grammatical simplification.

### 6.2.2 CT Quantifier Constructions

Another type of selecting subfunction of the modification function is quantification, or quantifier constructions. Croft defines these constructions as where a set's exact cardinality is not specified (forthcoming, p. 98). He presented four different types of quantifier constructions.

In CT Vague Enumeration Constructions, a set of countable items is selected, but the precise number is not specified. The word 𑜀𑜢𑜤𑜰𑜫 *tɕɔi'* 'many' fulfills this function, as in (47). It should be noted that there is no tone sandhi on the quantified ObjR in these constructions.

- (47) a. 有 人 𠵼 𠵼 耍 水  
u<sup>11</sup> naŋ<sup>55</sup> tɕɔi<sup>11</sup> na<sup>11</sup> suŋ<sup>52-11</sup> tɕui<sup>52</sup>  
have person many PROG play water  
‘There are many people playing in the water.’ [G3F1-1]
- b. 𠵼 伊 在 𠵼 厝. Ø 𠵼 厝 𠵼  
kiã<sup>52</sup> i<sup>33</sup> tɔ<sup>11</sup> t<sup>h</sup>iap<sup>5-2</sup> tɕ<sup>h</sup>u<sup>21</sup>. Ø t<sup>h</sup>iap<sup>5-2</sup> tɕ<sup>h</sup>u<sup>21</sup> tɕɔi<sup>11</sup>  
child 3SG PROG pack.up house. 3SG pack.up house many  
‘The child, she is packing up houses. [She] packs up many houses.’  
[G1F3-15]

In the CT Vague Enumeration Constructions in (47), the word order is ObjR + *tɕɔi*<sup>11</sup>. Comparing this usage to the literature, there are a few examples of what appears to be the same vague enumerator in Singapore Teochew (48a-b), Jambi Teochew (48c), and Jieyang (48d) varieties. Each of these examples shows pre-ObjR modification. Xu (2007) even showed how the Jieyang Vague Enumeration Construction can split the predicate-object compound idiomatic compounds, like *sue*<sup>2</sup> *ue*<sup>11</sup> ‘pray’ in (48d).

- (48) a. tsɿŋ<sup>33</sup> tɕɔi<sup>11</sup> naŋ<sup>55</sup> tsu<sup>11</sup> si<sup>11</sup> a<sup>33</sup> ne<sup>33</sup> k<sup>h</sup>uaŋ<sup>53</sup> hɔ̃<sup>53</sup>  
very much person SEQ COP like.that type PRT  
‘Many people are just like that.’ (Low, 2014, p. 49)
- b. ta<sup>2</sup> ta<sup>2</sup> ta<sup>2</sup> ta<sup>2</sup> kau<sup>53</sup> tseik<sup>2</sup> pua<sup>11</sup> t<sup>h</sup>ɔ̃<sup>53</sup>-tiɔ<sup>2</sup>  
ride ride ride ride until one half see-COMP.RES  
hi<sup>33</sup>kai<sup>11</sup>...u<sup>11</sup> sã<sup>33</sup>... tseŋ<sup>33</sup> tsui<sup>11</sup> kuɛ<sup>13</sup>tsi<sup>53</sup>  
DEM have three INT many fruit  
‘He keeps riding until halfway, he saw that...there were three...many fruits.’ (Yeo, 2011, p. 97)
- c. waa ziŋ<sup>33</sup> tɕɔi<sup>213</sup> ti<sup>11</sup>hẽ<sup>33</sup>...  
waa very many place  
‘Waa, many places...’ (Peng, 2012, p. 383)

- d. sueʔ<sup>2</sup> hoʔ<sup>2</sup> tsoi<sup>11</sup> ue<sup>11</sup>  
 say very many word  
 ‘pray for many things’ (Xu, 2007, p. 71)

Historically, we find examples of 𠵶 *tɕɔi*<sup>11</sup> + ObjR and ObjR + 𠵶 *tɕɔi*<sup>11</sup> in Fielde (1878), shown in (49). However, in (49a), 𠵶 *tɕɔi*<sup>11</sup> is acting as the predicate, which appears to be the case for all of the examples in Fielde (1878) which show ObjR + *tɕɔi*<sup>11</sup> order.<sup>66</sup>

- (49) a. 行李 𠵶  
 heŋ<sup>55</sup>li<sup>52</sup> tɕɔi<sup>11</sup>  
 baggage many  
 ‘much baggage’ (Fielde, 1878, p. 346)
- b. 𠵶 𠵶 人  
 many many person ...  
 tɕɔi<sup>11</sup> tɕɔi<sup>11</sup> naŋ<sup>55</sup> ...  
 ‘many people...’ (Fielde, 1878, p. 360)

In Khmer Vague Enumeration Constructions, modification occurs after the ObjR, as in (50). According to Sak-Humphry (2016, p. 224), in the examples in (50), the function of *klah* ‘some’ and *craən* ‘many’ is modification, unlike in (49a) that showed predication.

- (50) a. 𠵶 𠵶  
 koon klah  
 child some  
 ‘some children’ (Sak-Humphry, 2016, p. 224)
- b. 𠵶 𠵶  
 koon craən  
 child many  
 ‘many children’ (Sak-Humphry, 2016, p. 224)

<sup>66</sup> 𠵶 *tɕɔi*<sup>11</sup> ‘many’ is also frequently used for this function in Cambodian Teochew. See §7.3.

Given that Cambodian Teochew displayed the same ObjR + 𠵹 *tɕɔi<sup>11</sup>* ‘many’ word order in (47), there is potential that this order arose due to influence from Khmer.

However, there was one example of another CT Vague Enumeration Construction with the word 減 *kiam<sup>24</sup>* ‘few’ in (51) which shows ‘few’ + ObjR order. More examples are needed to confirm any patterns in terms of the behavior of this quantifier.

- (51) 減 卵  
kiam<sup>24</sup> nuŋ<sup>24</sup>  
few egg  
‘a few eggs’ [G1F1-PS]

Croft (forthcoming, p. 98) defined an amount term as a different type of selecting construction that identifies imprecise quantity for non-countable items. Some quantifiers can be used for both vague enumeration and amount. It appears the CT word 𠵹 *tɕɔi<sup>11</sup>* ‘many’ can be used in this respect, though there was only one instance where it was found in the present data set, in the compound in (52).

- (52) 𠵹谢  
tɕɔi<sup>11</sup>-sia<sup>11</sup>  
many-thanks  
‘Thank you.’

There are some examples in the literature from Singapore (53a) and Jieyang (53b-c) varieties. More data is needed to confirm the existence and behavior of this type of construction in Cambodian Teochew.

- (53) a. miŋ<sup>11</sup>ts<sup>h</sup>iŋ<sup>55</sup> pua<sup>ʔ2</sup> lo<sup>ʔ2</sup> lai<sup>11</sup> tsu<sup>33</sup> ie<sup>31</sup> buə<sup>35</sup> tsɔ̄i<sup>11</sup>  
 bed fall down come SEQ 3SG.POSS behind digit  
 le<sup>ʔ2</sup> tse<sup>11</sup> tsin<sup>33</sup> tsĩa<sup>53</sup> tua<sup>11</sup> le<sup>ʔ5</sup> lau<sup>11</sup> tsin<sup>33</sup> tsɔi<sup>11</sup> huə<sup>ʔ2</sup>  
 cut once very big cut flow very much blood  
 ‘When he fell down from the bed, he had a very big cut on his little toe and there was a lot of blood.’ (Low, 2014, p. 56)
- b. kio<sup>213-53</sup> k<sup>h</sup>a<sup>ʔ2-3</sup> tsoi<sup>11</sup> ts<sup>h</sup>ai<sup>213</sup> loŋ<sup>53-35</sup> tsiak<sup>5-2</sup> m<sup>35-21</sup> (tek<sup>2</sup>)-uaŋ<sup>55</sup>  
 order too much food all eat not (POT)-finish  
 ‘[We] ordered too much food and couldn’t finish it all.’  
 (Xu, 2007, p. 228)
- c. i<sup>33</sup> k<sup>h</sup>e<sup>ʔ2</sup> naŋ<sup>55-11</sup> t<sup>h</sup>au<sup>33</sup>-k<sup>h</sup>io<sup>ʔ5-2</sup> ho<sup>ʔ2</sup> tsoi<sup>11</sup> tsĩ<sup>55</sup>  
 3SG PASS person steal-take very much money  
 ‘He got a lot of his money stolen [by people].’ (Xu, 2007, p. 198)

Finally, per Croft (forthcoming, p. 99), in distributive quantifier constructions, the predicate applies to every member of the set by applying to each of the individuals in the set. There were some examples of CT Distributive Quantifier Constructions in my collected data. In (54a), the predicate (谢 *sia<sup>11</sup>* ‘thank’) is performed each year. Similarly, in (54b), the predicate (练 *lian<sup>11</sup>* ‘practice’) is done each afternoon from 2:00 to 5:00.

- (54) a. 每 年 谢 神 ...  
 mue<sup>52-24</sup> ni<sup>55</sup> sia<sup>11</sup> siŋ<sup>33</sup> ...  
 every year thank god ...  
 ‘[I] thank god every year...’ [G1M2-24]
- b. 母 日 个 下 午 在 只 个 两 点  
 mue<sup>52-24</sup> dzik<sup>5-2</sup> kai<sup>55-11</sup> e<sup>24-11</sup> kua<sup>21</sup> tɔ<sup>11</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> nɔ<sup>24</sup> tiam<sup>52</sup>  
 every day LW afternoon at this LW two o’clock  
 就 练 到 五 点  
 tɕiu<sup>11</sup> lian<sup>11</sup> kau<sup>21-53</sup> ŋɔu<sup>24</sup> tiam<sup>52</sup>  
 then practice until five o’clock  
 ‘Every afternoon [I] practiced from 2:00 to 5:00.’ [G1M2-24]

In the examples in (54), the CT Distributive Quantifier Constructions use the juxtaposition strategy where the word that expresses the quantification appears next to the modified ObjR. Speaker G2F1US also uses a strategy of reduplication to express the same function, as seen in (55).

- (55) a. “日日 我 恬 来 学 练 跳”  
 “dzik<sup>5-2</sup> dzik<sup>5</sup> ua<sup>52</sup> tiam<sup>11</sup> lai<sup>55-11</sup> ɔŋ<sup>5</sup> lian<sup>11</sup> t<sup>h</sup>iau<sup>21</sup>”  
 every.day 1SG always come learn practice jump  
 ‘Every day I always come to learn and practice jumping.’ [G2F1US-20]
- b. 伊 就 孛团 伊 问 但 “日日日日  
 i<sup>33</sup> tɕiu<sup>11</sup> nɔu<sup>11</sup>kiã<sup>52</sup> i<sup>33</sup> muŋ<sup>11</sup> tã<sup>21</sup> “dzik<sup>5-2</sup>dzik<sup>5</sup>dzik<sup>5-2</sup>dzik<sup>5</sup>  
 3SG then child 3SG ask say every.day  
 我 恬 行 过 睇 着 汝 .. 笑笑 ...”  
 ua<sup>52</sup> tiam<sup>11</sup> kiã<sup>55-11</sup> kue<sup>21</sup> t<sup>h</sup>ɔi<sup>52-24</sup> tiɔŋ<sup>2</sup> lu<sup>52</sup> .. tɕhio<sup>21-53</sup>tɕhio<sup>21</sup> ...”  
 1SG always walk across see RC 2SG .. smile RED ...  
 ‘Then the child asks “Everyday I always walk across and see you smiling...”’ [G2F1US-19]
- c. 人人 是 惊  
 naŋ<sup>55-11</sup>naŋ<sup>55</sup> si<sup>11</sup> kiã<sup>33</sup>  
 everyone COP fear  
 ‘Everyone is afraid.’ [G2F1US-18]

According to Xu (2007, p. 54), only reduplicated measure words and classifiers are used for this function in Jieyang.<sup>67</sup> Per Chao (1968/2011, p. 222), outside of measure words and classifiers in Chinese, only the reduplicated ObjR 人 ‘person’ provides a similar distributive

<sup>67</sup> A reduplicated classifier in other Sinitic languages, including Cantonese and Mandarin, also yields a distributive quantifier construction (Cheng & Sybesma, 2005, pp. 275-6).

function. More data is needed from more speakers to confirm the types of words that can be used in this function in Cambodian Teochew.

### 6.3 CT Object Modification Constructions

As previously mentioned, object modification can be of the situating or subcategorizing subfunctions of the modification function. Other terminology used to differentiate situating and subcategorizing object modification constructions is anchoring constructions and non-anchoring/typifying constructions, respectively.

Anchoring serves to uniquely identify the referent being modified (Croft, forthcoming, p. 129, citing Koptjevskaja-Tamm, 2002, 2003 and others). Anchoring prototypically denotes an individual, typically human, that is often expressed by a proper noun or contextual object (Croft, forthcoming, p. 130). CT Anchoring Constructions are covered in §6.3.1.

Conversely, typifying is characterized by having a modifier that is only type identifiable, that classifies a subclass of head objR, and that does not identify the head objR (Croft, forthcoming, p. 130, citing Koptjevskaja-Tamm, 2002, p. 154). CT Typifying Constructions are explored in §6.3.2.

#### 6.3.1 CT Anchoring Constructions (situating)

Croft (forthcoming) defines three different types of situating object modifiers which are illustrated with examples from Cambodian Teochew in Table 6.5. Each of these is an example of an anchoring construction because the modifiers identify the specific modified referent.

Information packaging function	Constructions		Semantic relations	Examples
situating	object modification construction	possession (alienable)	ownership relation	伊个房 i <sup>33</sup> kai <sup>55-11</sup> paŋ <sup>55</sup> ‘his room’
		possession (typically inalienable)	body part relation	伊脚 i <sup>33</sup> k <sup>h</sup> a <sup>33</sup> ‘his leg’
			kinship relation	伊妈妈 i <sup>33</sup> ma <sup>55-11</sup> ma <sup>55</sup> ‘her mother’
		locative modifier	figure-ground spatial relation	獐*鼠      □    土墙      顶畔 ŋiau <sup>11</sup> tɕ <sup>h</sup> u <sup>52</sup> na <sup>11</sup> t <sup>h</sup> ɔu <sup>11</sup> tɕ <sup>h</sup> ɔ <sup>55</sup> teŋ <sup>52-11</sup> paĩ <sup>55</sup> mouse      LOC wall      atop ‘mouse on top of a wall’

Table 6.5: CT Object Modification Constructions of the situating subfunction of the modification function, adapted from Croft (forthcoming, p. 95)

### 6.3.1.1 CT Possessive Modification Constructions

A distinction is made between alienable possession and inalienable possession because languages often use different constructions for these two possessive relationships. Alienable possession indicates ownership of an entity while inalienable possession might indicate a body-part relation, part-whole relation, or kinship relation. The distinction between what is considered alienable and inalienable is not universal but language-specific. It is the case that there is some distinction in CT Alienable Possessive Modification Constructions and CT Inalienable Possessive Constructions. We will show the criteria that separates these two constructions in Cambodian Teochew below.

CT Alienable Possessive Modification Constructions use the linker strategy with 个 *kaĩ<sup>55</sup>* ‘LW’, as in (56).

- (56) a. Felix 伊 口 伊 个 房 口 伊 个 褥  
*Felix* i<sup>33</sup> na<sup>11</sup> i<sup>33</sup> kai<sup>55-11</sup> paŋ<sup>55</sup> na<sup>11</sup> i<sup>33</sup> kai<sup>55-11</sup> dzək<sup>5</sup>  
 Felix 3SG LOC 3SG LW room LOC 3SG LW mattress  
 ‘Felix, he is in his room, in his crib.’ [G3F1-2]
- b. 马打 就 互 Ø 伊 个 衫裤  
 ma<sup>33</sup>ta<sup>24</sup> tɕiu<sup>11</sup> həu<sup>11</sup> Ø i<sup>33</sup> kai<sup>55-11</sup> sã<sup>33</sup>k<sup>h</sup>əu<sup>21</sup>  
 police then give 3SG 3SG LW clothes  
 ‘[The] police then give [him] his clothes.’ [G3F2-4]
- c. 伊人 个 花园  
 i<sup>33</sup>naŋ<sup>11</sup> kai<sup>55-11</sup> hue<sup>33</sup>huŋ<sup>55</sup>  
 3PL LW flower-garden  
 ‘...their garden.’ [G1F4-16]
- d. 猫 个 故事  
 ŋiau<sup>33</sup> kai<sup>55-11</sup> ku<sup>52</sup>su<sup>11</sup>  
 cat LW story  
 ‘The Cat Story.’ [G3F1-3]

In CT Alienable Possessive Modification Constructions, the linker undergoes tone sandhi while the modifying object does not. This is illustrated in (57).

- (57) a. 猪 个 厝  
 tu<sup>55</sup> kai<sup>55-11</sup> tɕ<sup>h</sup>u<sup>21</sup>  
 pig LW house  
 ‘[the] pig’s house’ [G2M1-5]
- b. 蒙古 个 厝 ...蒙古人 个 厝  
 mŋ<sup>11</sup>kəu<sup>52</sup> kai<sup>55-11</sup> tɕ<sup>h</sup>u<sup>21</sup> ... mŋ<sup>11</sup>kəu<sup>52-24</sup>naŋ<sup>55</sup> kai<sup>55-11</sup> tɕ<sup>h</sup>u<sup>21</sup>  
 Mongolia LW house ... Mongolia person LW house  
 ‘A Mongolian house... a Mongolian person’s house’ [G1M1-7]
- c. 伊 欲 去 觅 白兔 个 内  
 i<sup>33</sup> ai<sup>21-53</sup> k<sup>h</sup>u<sup>21</sup> tɕ<sup>h</sup>ue<sup>11</sup> pe<sup>ʔ5-2</sup>t<sup>h</sup>əu<sup>21</sup> kai<sup>55-11</sup> lai<sup>24</sup>  
 3SG want go find rabbit LW home  
 ‘He wants to go find the rabbit’s house.’ [G2F1US-18]

- d. 就 伊 个.. 鸟 个 媛 就 来  
 tɕiu<sup>11</sup> i<sup>33</sup> kai<sup>55-11</sup>.. tɕiau<sup>52</sup> kai<sup>55-11</sup> ai<sup>55</sup> tɕiu<sup>11</sup> lai<sup>55</sup>  
 then 3sg LW.. bird LW mother then come  
 ‘Then its.. the bird’s mother then comes.’ [G3F1-3]

CT Inalienable Possession Constructions are also able to use the linker strategy. This is seen for body parts in (58) and for kinship terms in (59).

- (58) a. 伊 来.. 滚\* 口.. 嫖妘 个 脚  
 i<sup>33</sup> lai<sup>55-11</sup>.. nua<sup>52</sup> na<sup>11</sup>.. tɕa<sup>33</sup> bɔu<sup>33</sup> kai<sup>55-11</sup> k<sup>h</sup>a<sup>33</sup>  
 3SG come.. roll LOC.. woman POSS leg  
 ‘He comes and rolls at the woman’s leg.’ [G3F1-3]
- b. 伊 个 手 在 只 个 丈夫 个 手  
 i<sup>33</sup> kai<sup>55-11</sup> tɕ<sup>h</sup>iu<sup>52</sup> tɔ<sup>11</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> ta<sup>33</sup> pɔu<sup>33</sup> kai<sup>55-11</sup> tɕ<sup>h</sup>iu<sup>52</sup>  
 3SG LW hand LOC this LW man LW hand  
 ‘Her hand is on this man’s hand.’ [G3F2-4]
- c. 鱼 个 目  
 hu<sup>55</sup> kai<sup>55-11</sup> mak<sup>5</sup>  
 fish LW eye  
 ‘fish’s eye’ [G1F2-13]
- d. 丈夫 个 尻仓  
 ta<sup>33</sup> pɔu<sup>33</sup> naŋ<sup>55</sup> kai<sup>55-11</sup> k<sup>h</sup>a<sup>33</sup> tɕ<sup>h</sup>uŋ<sup>33</sup>  
 man LW butt  
 ‘man’s butt’ [G2F1US-PS]
- (59) a. Ø 行 觅 伊 个 团  
 Ø kia<sup>55-11</sup> tɕ<sup>h</sup>ue<sup>11</sup> i<sup>33</sup> kai<sup>55-11</sup> kiã<sup>52</sup>  
 3SG leave look for 3SG LW child  
 ‘[She] leaves to look for her child.’ [G3F1-2]

- b. 伊 个 妯 合 伊 个 丈人 合 伊 个 团...  
 i<sup>33</sup> kai<sup>55-11</sup> bəu<sup>52</sup> ka<sup>ʔ2-5</sup> i<sup>33</sup> kai<sup>55-11</sup> ti<sup>ʃ</sup><sup>11</sup> naŋ<sup>55</sup> ka<sup>ʔ2-5</sup> i<sup>33</sup> kai<sup>55-11</sup> kiã<sup>52</sup> ...  
 3SG LW wife with 3SG LW father.of.wife with 3SG LW kid...  
 ‘His wife and his wife’s father and his kid...’ [G3F2-4]
- c. 只 个 嫜 个 翁人 ...  
 tɕei<sup>52</sup> kai<sup>55-11</sup> tɕa<sup>33</sup> bəu<sup>52</sup> kai<sup>55-11</sup> aŋ<sup>33</sup> naŋ<sup>55</sup> ...  
 this LW woman LW husband ...  
 ‘This woman’s husband...’ [G3F2-4]
- d. Ø 行 搭 伊 个 孙  
 Ø kia<sup>55-11</sup> ta<sup>ʔ5</sup> i<sup>33</sup> kai<sup>55-11</sup> suŋ<sup>33</sup>  
 3SG walk together 3SG LW grandchild  
 ‘[He] walks together with his grandchild.’ [G3F1-1]

At the same time, CT Inalienable Possessive Constructions can also use the juxtaposition strategy where the possessor and the head ObjR appear next to each other with no additional morpheme. The examples given in (60) showcase CT Inalienable Possessive Constructions featuring body parts modified using the juxtaposition strategy, while the examples in (61) show kinship relations.

- (60) a. 只个 Ø 坐 在 伊 背脊  
 tɕei<sup>55</sup> Ø tɕə<sup>24-11</sup> tɔ<sup>11</sup> i<sup>33</sup> pa<sup>33</sup> tɕia<sup>ʔ2</sup>  
 PDCO 3SG sit LOC 3SG.POSS back  
 ‘[In] this (one [the picture]), [he] sits on its back.’ [G1F1-9]
- b. □ 只个 伊 揽 伊 脚  
 ɔ<sup>11</sup> tɕei<sup>55</sup> i<sup>33</sup> lam<sup>52</sup> i<sup>33</sup> k<sup>h</sup>a<sup>33</sup>  
 INTJ PDCO 3SG hug 3SG.POSS leg  
 ‘[In] this (one [the picture]), he hugs his leg.’ [G1F1-9]
- c. 只个 Ø 睇 电视, 伊 脚 伸 落 去  
 tɕei<sup>55</sup> Ø t<sup>h</sup>ɔ<sup>ʔ52-24</sup> tiaŋ<sup>11</sup> si<sup>24</sup>, i<sup>33</sup> ka<sup>33</sup> tɕ<sup>h</sup>uŋ<sup>33</sup> lɔ<sup>ʔ5</sup> ku<sup>21</sup>  
 PDCO 3SG watch TV, 3SG.POSS feet stretch descend go  
 ‘[In] this (one [the picture]), he watches TV and stretches out his legs.’  
 [G1F1-8]

- d. ... 就 隻 狗.. 口 在 伊 头 顶  
 ... tɕiu<sup>11</sup> tɕiaŋ<sup>2-5</sup> kau<sup>52</sup>.. p<sup>h</sup>ɔŋ<sup>2</sup> tɔ<sup>11</sup> i<sup>33</sup> t<sup>h</sup>au<sup>55-11</sup> teŋ<sup>52</sup>  
 ... then CLF.IK dog.. hang LOC 3SG.POSS head atop  
 ‘...then the dog is hang on top of his head.’ [G1F1-11]
- e. 好 伊 行 去 在 伊 尻仓 後尾  
 hɔ<sup>52</sup> i<sup>33</sup> kiã<sup>55-11</sup> k<sup>h</sup>u<sup>21</sup> tɔ<sup>11</sup> i<sup>33</sup> k<sup>h</sup>a<sup>33</sup> tɕ<sup>h</sup>uŋ<sup>33</sup> aɔ<sup>11</sup>bue<sup>52</sup>  
 good 3SG walk go LOC 3SG.POSS butt behind  
 ‘The he walks behind his butt.’ [G2F1US-18]
- (61) a. 伊 合 伊 团 但, 合 伊 妘 但  
 i<sup>33</sup> kaŋ<sup>2-5</sup> i<sup>33</sup> kiã<sup>52</sup> tã<sup>21</sup>, kaŋ<sup>2-5</sup> i<sup>33</sup> bɔu<sup>52</sup> tã<sup>21</sup>  
 3SG with 3SG.POSS child talk, with 3SG.POSS wife talk  
 ‘He talks with his child, talks with his wife.’ [G3F3-4]
- a. 我 爸 在 ...  
 ua<sup>52</sup> pa<sup>33</sup> tɔ<sup>24</sup> ...  
1SG.POSS father live ...  
 ‘When my father was alive...’ [G1F6-24]
- c. 伊 妈妈  
 i<sup>33</sup> ma<sup>55-11</sup> ma<sup>55</sup>  
3SG-POSS mother  
 ‘his mother’ [G1F1-10, n = 4; G1F4-16, n = 6]
- d. 伊 媛  
 i<sup>33</sup> ai<sup>55</sup>  
3SG-POSS mother  
 ‘his mother’ [G1F2-13, n = 6; G2M1-6, n = 12]

Kinship relations extend beyond strictly familial relations to include other relationships, such as those in (62), which are further examples of CT Inalienable Possessive Constructions using the juxtaposition strategy.

- (62) a. ... 睇 着 伊 主人 在 食饭  
 ... thoŋ<sup>52-24</sup> tiŋ<sup>2</sup> i<sup>33</sup> tɕu<sup>52-24</sup> dʒiŋ<sup>55</sup> tɔ<sup>11</sup> tɕiaŋ<sup>55-2</sup>-puŋ<sup>11</sup>  
 ... see RC 3SG.POSS master PROG eat.rice  
 ‘[The cat] sees its master eating.’ [G1F1-9]
- b. Ø 去 食 酒 合 伊 朋友  
 Ø khu<sup>21-53</sup> tɕiaŋ<sup>55-2</sup> tɕiu<sup>52</sup> kaŋ<sup>2-5</sup> i<sup>33</sup> p<sup>h</sup>ej<sup>55-11</sup> iu<sup>52</sup>  
 3sg go drink alcohol with 3SG.POSS friend  
 ‘[He] goes to drink alcohol with his friend(s).’ [G3F3-4]

So far we have shown how CT Alienable Possessive Constructions require the use of the linker 个 *kaŋ<sup>55</sup>* ‘LW’, while CT Inalienable Constructions are able to use both the linker and the juxtaposition strategies. This difference follows the universal outlined in Croft (forthcoming, p. 139), that inalienable possession will be expressed with the same number or fewer morphemes than alienable possession. Citing Haiman (1983), Croft says this is due to a closer “conceptual distance” that manifests in an iconic “linguistic distance” (forthcoming, p. 139).

In Singapore Teochew, Yeo (2011, p. 54) argued that the linker *kaŋ<sup>11</sup>* is optional only when the possessor is a contextual object and the possessed object is a familial term or body part. This is seen in her examples below, where (63a) and (63a) show optional *kaŋ<sup>11</sup>* with *i<sup>33</sup>* ‘3SG’ as a possessor, while the linker strategy is used when the modifier is a Proper ObjR in (63c).

- (63) a. i<sup>33</sup>(kaŋ<sup>11</sup>) bou<sup>53</sup> tɔ<sup>11</sup> laŋ<sup>53</sup>  
 3SG(-GEN) wife LOC home  
 ‘His wife is at home.’ (Yeo, 2011, p. 54)
- b. i<sup>33</sup>(kaŋ<sup>11</sup>) ts<sup>h</sup>iu<sup>53</sup> tsem<sup>33</sup>-t<sup>h</sup>ia<sup>13</sup>  
 3SG(-GEN) hand INT-be.pain.  
 ‘His hand hurts a lot.’ (Yeo, 2011, p. 55)

- c. Mary-kai<sup>11</sup> k<sup>h</sup>a<sup>33</sup> t<sup>h</sup>ia<sup>13</sup> bɔɪ<sup>11</sup> kia<sup>33</sup>  
 Mary-GEN leg be.pain NEG walk  
 ‘Mary’s leg hurts (so) she cannot walk.’ (E) (Yeo, 2011, p. 34)

This restriction might also be present in Cambodian Teochew. All of the examples of the juxtaposition strategy in (60-62) and elsewhere in the texts are cases where the Contextual Object 伊 *i*<sup>33</sup> ‘3SG’ is the modifier, with the exception of (61b) with Contextual Obj 我 *ua*<sup>52</sup> ‘1SG’. Conversely, for each instance in the texts where a Common ObjR is the modifier, including the examples in (57) and (58), the linker strategy is used. However, there were some examples in the picture stimuli data from speaker G1F4 where the juxtaposition strategy was used for body parts of animals. Compare the pairs of utterances in (64) where G1F4 produced the CT Inalienable Possessive Constructions with the juxtaposition strategy on the left, while speaker G1F3 produced the CT Inalienable Possessive Constructions with the linker strategy on the right, for the same prompt.

- |      |    |   |  |
|------|----|---|--|
| (64) | a. | 牛 脚<br>gu <sup>55-11</sup> k <sup>h</sup> a <sup>33</sup><br>cow foot<br>‘cow’s foot’ [G1F4-PS] | 伊 个 脚<br>i <sup>33</sup> kai <sup>55-11</sup> k <sup>h</sup> a <sup>3</sup><br>3SG LW foot<br>‘its foot’ [G1F3-PS] |
|      | b. | 牛 尾<br>gu <sup>55-11</sup> bue <sup>52</sup><br>cow tail<br>‘cow’s tail’ [G1F4-PS]              | 伊 个 尾<br>i <sup>33</sup> kai <sup>55-11</sup> bue <sup>52</sup><br>3SG LW tail<br>‘its tail’ [G1F3-PS]             |
|      | c. | 牛 角<br>gu <sup>55-11</sup> kak <sup>4</sup><br>cow horn<br>‘cow’s horn’ [G1F4-PS]               | 伊 个 角<br>i <sup>33</sup> kai <sup>55-11</sup> kak <sup>4</sup><br>3SG LW horn<br>‘its horn’ [G1F3-PS]              |

However, it is possible that the utterances on the left in (64) are actually CT Typifying Constructions, where the modifier 牛 *gu*<sup>55</sup> ‘cow’ is used to describe a certain type of ‘foot’, ‘tail’, and ‘horn’, as opposed to situating/anchoring one particular instance of those body parts. If they are indeed Anchoring Constructions, then it seems that for at least some speakers in some contexts, the juxtaposition strategy is used with Inalienable Possessive Constructions with modifiers that are not contextual objects.

We could be seeing evidence of the Extended Animacy Hierarchy here. In Singapore Teochew, it appears that the type of referents that can go in the modifier slot in Inalienable Possessive Constructions with the juxtaposition strategy is restricted to Contextual Objects, or things that are high on the EAH. In Cambodian Teochew, if non-human animates are allowed as the possessor, we would posit that CT Inalienable Possessive Constructions with human modifiers should also be allowed to use the juxtaposition strategy. More research needs to be done in this area to verify and expand on these findings.

In addition to the strategies already explored, the literature on other Teochew varieties found other strategies for expressing possessive constructions that were not found in the current data set on Cambodian Teochew, including the classifier strategy and the special form strategy.

Xu (2007, p. 80) showed how Jieyang Teochew Possessive Constructions can use the classifier strategy with a contextual object as the possessor (65a-e), or with a common object (65f).<sup>68</sup>

---

<sup>68</sup> I have added characters to examples in (65) from Xu (2007) and adjusted the glossing of the classifiers.

- (65) a. 我 件 衫  
ua<sup>53</sup> kiã<sup>35-21</sup> sã<sup>33</sup>  
1SG CLF.ITEM blouse/shirt  
‘my blouse/shirt’ (Xu, 2007, p. 80)
- b. 我 支 笔  
ua<sup>53</sup> ki<sup>33</sup> peŋ<sup>2</sup>  
1SG CLF.STICK pen  
‘my pen’ (Xu, 2007, p. 113)
- c. 伊 隻 狗  
i<sup>33</sup> tsiaŋ<sup>2-3</sup> kau<sup>53</sup>  
3SG CLF.ANIMAL dog  
‘his dog’ (Xu, 2007, p. 113)
- d. ... 伊 张 脚车  
... i<sup>33</sup> tiõ<sup>33</sup> k<sup>h</sup>a<sup>33</sup>-ts<sup>h</sup>ia<sup>33</sup>  
... 3SG CLF.FLAT foot-vehicle  
‘... his bike’ (Xu, 2007, p. 235)
- e. 伊 双 目 大大 双  
i<sup>33</sup> saŋ<sup>33</sup> mak<sup>5</sup> tua<sup>11</sup>-tua<sup>11</sup> saŋ<sup>33</sup>  
3SG pair eye big-big pair  
‘Her eyes are very big.’ (Xu, 2007, p. 111)
- e. 老师 本 书  
lau<sup>53-35</sup>su<sup>33</sup> puŋ<sup>53-35</sup> tsu<sup>33</sup>  
teacher CLF.BOOK book  
‘the teacher’s book(s)’ (Xu, 2007, p. 113)

Similarly, there was some evidence from X. Chen (2003) that JBT Possessive Constructions use the classifier strategy, in (66).

- (66) ... 我 廉 柑  
... ua<sup>53</sup> liam<sup>213</sup> kã<sup>33</sup>  
... 1SG CLF.SLICE orange  
‘...my slice of orange.’ (X. Chen, 2003, pp. 381-2)

No evidence was found for the use of the classifier strategy in CT Possessive Constructions. This likely relates to the patterns we saw earlier in the gravitation towards the use of the linker strategy for Enumeration Modification Constructions. Since the linker strategy is a perfectly acceptable form to use for the possessive function, it is much simpler to use *kar*<sup>55</sup> for all ObjRs rather than using the classifier strategy which would change depending on the modified ObjR. The lack of any evidence for this strategy for CT Possessive Constructions for the speakers here represents another example of the simplification of the grammar of Cambodian Teochew, an occurrence typical of contact languages.

Xu (2007, p. 74) argued that there are fused forms used in Jieyang Inalienable Possessive Constructions with kinship terms, where *ua*<sup>53</sup> + *kar*<sup>55</sup> fuse to form *uã*<sup>33</sup> ‘my/our’ and *lu*<sup>53</sup> + *kar*<sup>55</sup> fuse to form *nia*<sup>33</sup> ‘your (SG/PL)’. In Croft’s terminology, this would be the use of the special form strategy because it only occurs in this one construction (Croft, forthcoming, p. 124).<sup>69</sup> The special form strategy is not allowed with other terms in Jieyang such as body parts or common ObjRs, as in (68) from Xu (2007, p. 79).

- (67) a.    *uã*<sup>33</sup>            *pa*<sup>55</sup>  
           1SG.GEN        father  
           ‘my father’
- b.    *nia*<sup>33</sup>            *pa*<sup>55</sup>  
           2SG.GEN        father  
           ‘your father’
- c.    *uã*<sup>33</sup>            *pa*<sup>55</sup>    *oi*<sup>53-35</sup>    *uã*<sup>33</sup>            *ma*<sup>55</sup>  
           1SG.GEN        father    short    1SG.GEN        mother  
           ‘My father is shorter than my mother.’ (Xu, 2007, p. 272)

---

<sup>69</sup> The special form strategy is also found in Hakka for inalienable possession (Chappell, 2001, pp. 351-2).

- (68) a. \*uã<sup>33</sup> mak<sup>2</sup>  
 1SG.GEN eye  
 ‘\*my eye’
- b. \*nia<sup>33</sup> tsu<sup>33</sup>  
 2SG/PL.GEN book  
 ‘\*your book’

According to Xu (2007), special forms in Jieyang exist only for first and second person possessors. In Singapore Teochew, Low (2014, p. 19) similarly found the special form strategy used for possession of kinship relations, and it extends to the third person as well, as shown in Table 6.6.

	Type 1	Type 2
1 <sup>st</sup> person (SG)	ŋa <sup>33</sup>	uŋ <sup>35</sup> ~ ɿ <sup>35</sup>
2 <sup>nd</sup> person (SG/PL)	nia <sup>33</sup>	niŋ <sup>35</sup>
3 <sup>rd</sup> person (SG, PL)	ia <sup>33</sup>	iŋ <sup>35</sup>

Table 6.6: Singapore Teochew Possessive Contextual Object Modifiers used for kinship terms (Low, 2014, p. 19)

No evidence was found for any of these special forms in Cambodian Teochew. However, given the extremely few tokens of first and second person Contextual Objects in the data set, it is certainly possible that they are used. Future research should focus on gathering more data such as personal narratives that would be conducive to these types of constructions.

### 6.3.1.2 CT Locative Modification Constructions

In locative modification constructions, such as ‘the bicycle in the garage’ in English, a ground object (garage) modifies the figure (bicycle), thus expressing a figure-ground spatial relation (Croft, forthcoming, p. 101).

Only a few CT Locative Modification Constructions were found in the collected data, provided in (69). Examples (69a-c) show perception events with the predicate 睇 *tʰɔ̃<sup>52</sup>* ‘see’, while (69d) shows a CT Presentational Construction. These are all CT Anchoring Constructions because the locative phrase acts to particularize the ObjR it is modifying. In (69d), there is not just any ‘bird’, there is the particular bird that is on top of a tree. CT Locative Modification Constructions use relational strategies in which an additional morpheme or morphemes is/are used to showcase the relationship between the figure ObjR and the ground ObjR. In the examples below, (69a) and (69d) use a circumpositional strategy with two adpositions, before and after the ground ObjR. In (69b) and (69c), a prepositional strategy is used and the relational morpheme appears before the ground ObjR.

- (69) a. 伊  □  行  就  睇见  
*i<sup>33</sup>    na<sup>11</sup>    kiã<sup>55</sup>    tɕiu<sup>11</sup>    tʰɔ̃<sup>52-11</sup>kiã<sup>21-53</sup>*  
 3SG    PROG    walk    then    see  
 獠\*鼠                    □    土墙                    顶畔  
*ŋiau<sup>11</sup>tɕ<sup>h</sup>u<sup>52</sup>                    na<sup>11</sup>    tʰɔu<sup>11</sup>tɕ<sup>h</sup>io<sup>55</sup>    ten<sup>52-11</sup>paĩ<sup>55</sup>*  
 mouse                    LOC    wall                    atop  
 ‘He is walking then [he] sees a mouse on top of a wall.’ [G3F1-3]
- b. 伊 睇见                    只 撮                    鸟团                    □ 鸟宿  
*i<sup>33</sup>    tʰɔ̃<sup>52-11</sup>kiã<sup>21-53</sup>                    tɕi<sup>52</sup>    tɕ<sup>h</sup>ɔ̃<sup>2-5</sup>                    tɕiau<sup>52-11</sup>-kiã<sup>52</sup>    na<sup>11</sup>    tɕiau<sup>52-11</sup>-siu<sup>21</sup>*  
 3SG    see                    this    CLF.GROUP    bird-baby                    LOC    bird-nest  
 ‘He sees these baby birds in a bird nest.’ [G3F1-3]
- c. 老姐..                    伊  睇  报纸                    □..    □    沙  
*lau<sup>24-11</sup>tɕe<sup>52</sup>..    i<sup>33</sup>    thɔi<sup>52-11</sup>    pɔ<sup>53</sup>tɕua<sup>52</sup>                    na<sup>11</sup>..    na<sup>11</sup>    sua<sup>33</sup>*  
 girl..                    3SG    see    newspaper                    LOC..    LOC    sand  
 ‘[The] girl, she sees the newspaper in the sand.’ [G3F1-1]
- d. ... 有  鸟    □    顶  树    顶畔  
 ...    u<sup>24</sup>    tɕiau<sup>52</sup>    na<sup>11</sup>    ten<sup>52-11</sup>    tɕ<sup>h</sup>iu<sup>11</sup>    ten<sup>52-11</sup>paĩ<sup>55</sup>  
 ...    have    bird    LOC    atop    tree    atop  
 ‘...there is a bird on top of a tree.’ [G3F1-3]

### 6.3.2 CT Typifying Constructions (subcategorizing)

Non-anchoring/typifying constructions subcategorize referents. In this way, they resemble property modification constructions, except here the modifier is an object.

Croft (forthcoming, p. 131) presented a continuum of constructions that illustrates how conventional an object modification construction is for naming a referent. I have adapted this continuum in (70).

(70) anchoring      →      typifying      →      dual object lexeme      →      common  
construction      construction      construction      objR

We've already covered CT Anchoring Constructions in §6.3.1. These terms are not close to being conventional terms for referents because they are anchored by one unique object. The three other types of constructions in (70) are all examples of typifying constructions, but they vary in terms of how conventionalized the object-object combination is for calling out a particular referent. The determination as to where the referent falls on the continuum varies based on the language and speech community. As explained by Croft (forthcoming), “[o]nce the [object-object] combination develops a unitary conventional meaning, it comes to resemble a unitary referent expression, and through lexicalization becomes a single lexical item.” (p. 141).

For the least conventionalized typifying constructions, the object modifier acts like a property in that it serves the purpose of adding additional information about the referent. This is seen in the examples of CT Typifying Constructions in (71).

- (71) a. 苹果 园  
 p<sup>h</sup>ej<sup>33</sup>kuẽ<sup>52-24</sup> huŋ<sup>55</sup>  
 apple garden  
 ‘apple garden’ [G1M1-7]
- b. 苹果 树  
 p<sup>h</sup>ej<sup>33</sup>kuẽ<sup>52-24</sup> tɕ<sup>h</sup>iu<sup>11</sup>  
 apple tree  
 ‘apple tree’ [G1M1-7]
- c. 獠\*鼠 空  
 ŋiau<sup>24</sup>tɕ<sup>h</sup>u<sup>52-24</sup> k<sup>h</sup>aŋ<sup>33</sup>  
 mouse hole  
 ‘mouse hole’ [G1F1-11]
- d. 衫裤 个 篮  
 sã<sup>33</sup>-k<sup>h</sup>ɔu<sup>21</sup> kai<sup>55-11</sup> na<sup>55</sup>  
 clothes POSS basket  
 ‘clothes basket’ [G3F1-2]

The examples in (71a-c) show CT Typifying Constructions with the juxtaposition strategy and (71d) uses the linker strategy. However, the strategy used does not dictate the function. Thus, juxtaposition, compounding, and linker strategies are not reserved for any one type of construction in (70). There are some instances where the same concept can be expressed with different strategies, depending on the speaker or context, such as in the CT Typifying Constructions in (72), where the first example uses the juxtaposition strategy and the second, the linker strategy.

- (72) a. 潮州 文化  
 tiɔ<sup>11</sup>tɕiu<sup>33</sup> buŋ<sup>11</sup>hue<sup>21</sup>  
 Teochew culture  
 ‘Teochew culture’ [G1F7-24]

- b.      潮州            个      文化  
           tio<sup>11</sup>tɕiu<sup>33</sup>      kai<sup>55-11</sup> buŋ<sup>11</sup>hue<sup>21</sup>  
           Teochew      LW      culture  
           ‘culture of Teochew [G1F6-24]

Further into the lexicalization process comes dual object lexeme constructions, a term adapted from Pepper (2020). Dual object lexeme constructions function “to provide generic names for complex concepts utilizing the names of two existing concepts between which there is an implicit, but unstated, relation” (p. 7). Thus, dual object lexeme constructions are closer to being lexicalized because they refer to a single concept.

My consultants were highly skilled in using the language and vocabulary that they knew in order to convey their message. At times this required what seemed to be the coining of new terms, which were often CT Dual Object Lexeme Constructions, such as the examples below which show compounding (73a-b), juxtaposition (73c-d), and linker strategies (73e).<sup>70</sup>

- (73) a.      位沙  
           ui<sup>11</sup>-sua<sup>33</sup>  
           place-sand  
           ‘desert’ [G3F1-1]
- b.      猫鸟  
           ŋiau<sup>33</sup>-tɕiau<sup>52</sup>  
           cat-bird  
           ‘owl’ [G1F4-16]

---

<sup>70</sup> An overall reduction in unique vocabulary terms would not be surprising in a context like that of Cambodian Teochew where the domains the language is used are limited and thus the speakers’ exposure to a wide range of vocabulary items might similarly be limited. This would then necessitate some lexical innovations (including language borrowing. See McFarland (2021)).

- c.     警察           厝  
*jǐngchá*        tɕ<sup>h</sup>u<sup>21</sup>  
 police           building  
 ‘prison’ [G3F2, G3F3-4]
- d.     电车           地方  
 tian<sup>11</sup>tɕ<sup>h</sup>ia<sup>33</sup>   ti<sup>11</sup>huŋ<sup>33</sup>  
 car              place  
 ‘parking lot’ [G1F1-8]
- c.     马    个    交椅  
 be<sup>52</sup>   kai<sup>55-11</sup>   kau<sup>33</sup>ĩ<sup>52</sup>  
 horse   LW    chair  
 ‘saddle’ [G3F1-1]

The right-most stage of Croft’s continuum in (70) is for generally accepted terms for a single concept. Sinitic languages in general are known for their high degree of “compounding”, or the use of two or more concepts to create a single concept.<sup>71</sup> In Pepper’s (2020, p. 124) study, dual object lexemes made up 77% (105/177) of Mandarin’s total vocabulary size, of the meanings the researchers were investigating. Dual object lexemes, as well as the more lexicalized common objects, are similarly abundant in Cambodian Teochew.

The examples in (74) show different types of 肉 *ba<sup>2</sup>* ‘meat’. They are all examples of CT Typifying Constructions in that the Object Modifier acts to subcategorize the meat. At the same time, they are examples of CT Common ObjRs because their usages are highly conventionalized. They are found in dictionaries and are the generally accepted terms for these concepts.

---

<sup>71</sup> See Arcodia (2007) for an analysis of “compounding” in Mandarin Chinese.

- |      |    |                                       |    |                                       |    |                                    |
|------|----|---------------------------------------|----|---------------------------------------|----|------------------------------------|
| (74) | a. | 牛肉                                    | b. | 羊肉                                    | c. | 猪肉                                 |
|      |    | gu <sup>55-11</sup> -baŋ <sup>2</sup> |    | iŋ <sup>55-11</sup> -baŋ <sup>2</sup> |    | tu <sup>33</sup> -baŋ <sup>2</sup> |
|      |    | cow-meat                              |    | sheep-meat                            |    | pig-meat                           |
|      |    | ‘beef’                                |    | ‘mutton’                              |    | ‘pork’                             |

Other examples of lexicalized CT Object Modification Constructions are given in (75). Given the frequent occurrence of this type of construction, these examples are only a snapshot of the possibilities which can include anything from persons, to artifacts, locations, and more.

- |      |    |  |    |  |
|------|----|--|----|--|
| (75) | a. | 头家   | b. | 鱼美人  |
|      |    | t <sup>h</sup> au <sup>55-11</sup> ke <sup>33</sup>    |    | hu <sup>55-11</sup> -mui <sup>52-24</sup> -dziŋ <sup>55</sup>    |
|      |    | head-house   |    | fish-pretty-person   |
|      |    | ‘boss, owner’ [G1F1-9]                                 |    | ‘mermaid’ [G1M1-7]   |
|      | c. | 电脑   | d. | 脚车   |
|      |    | tiaŋ <sup>55-11</sup> -nau <sup>52</sup>               |    | k <sup>h</sup> a <sup>33</sup> -tɕ <sup>h</sup> ia <sup>33</sup> |
|      |    | electricity-brain                                      |    | foot-vehicle   |
|      |    | ‘computer’ [G1F1-12]                                   |    | ‘bicycle’ [G1F1-10, G1M1-7, G1F4-16]                             |
|      | e. | 潮州话  | f. | 鸟宿   |
|      |    | tiə <sup>11</sup> tɕiu <sup>33</sup> -ue <sup>11</sup> |    | tɕiau <sup>52-11</sup> -siu <sup>21</sup>                        |
|      |    | Teochew-language                                       |    | bird-nest  |
|      |    | ‘Teochew’ [G1F6-24]                                    |    | ‘bird nest’ [G1F1-9, G3F1-3]                                     |
|      | g. | 餐馆   |    | 海舷   |
|      |    | tɕ <sup>h</sup> aŋ <sup>33</sup> kuaŋ <sup>52</sup>    |    | hai <sup>52-11/24</sup> ki <sup>55</sup>                         |
|      |    | food-building  |    | ocean-edge   |
|      |    | ‘restaurant’ [G1F4-16]                                 |    | ‘beach, shore’ [G3F1-1, G3F2-4, G1F3-15]                         |
|      | f. | 花园   |    |  |
|      |    | hue <sup>33</sup> -huŋ <sup>55</sup>                   |    |  |
|      |    | flower-garden  |    |  |
|      |    | ‘garden’ [G1F1-10, G1F3-15, G1F4-16, G2M1-5]           |    |  |

#### 6.4 CT Action Modification Constructions

Actions/events can act as a modifier to a referent. In action modification constructions, there will be (at least) two events, with one action/event modifying the main action/event. Croft

differentiates between two types of action/event modifiers. For him, “relative clause” constructions necessarily have a shared participant in both events (Croft, forthcoming, p. 544). In the English Action Modification Construction in (76) ‘the cake’ is a participant in both the ‘eating’ and the ‘baking’ events.

(76) John ate the cake [that she baked on Sunday].

Conversely in Croft’s (forthcoming, p. 556) “noun complement” constructions, the modified objR is not necessarily a participant of the modifying action clause. In the Japanese Action Modification Construction in (77), *zizitu* ‘fact’ is not an argument of the clause ‘the student bought the book’.

(77) [gakusei        ga        hon        o        katta] zizitu  
student        NOM    book    ACC    bought fact  
‘the fact that the student bought the book’ (Croft, forthcoming, p. 556)

Despite this difference, the function of the constructions in (76) and (77) remains the same in that the bracketed clauses are both actions modifying a referent. Thus, we follow LaPolla (2017, p. 101) in his findings on Mandarin to simply call all constructions of this type CT Action Modification Constructions.

Dryer (2013) divides languages depending on the order of the head objR and the action modifier clause. Here we’ll call the word order with head objR + action modifier clause a post-objR strategy and that with action modifier clause + head objR a pre-objR strategy.<sup>72</sup>

---

<sup>72</sup> Dryer (2013) abbreviates these orderings as NRel and RelN, while Croft (forthcoming, p. 546) calls them the “postnominal” and “prenominal” strategies, respectively.

Sinitic languages have been known to show the pre-objR strategy, such as in the Jieyang (78a) and Singapore (79b) Teochew Action Modification Constructions. These examples are also showing the use of the linker strategy, with 个 *kai*<sup>55</sup>.

- (78) a. ua<sup>53</sup> hi<sup>53-35</sup>huã<sup>33</sup> [i<sup>33</sup> li<sup>53-35</sup>lo<sup>5</sup>] kai<sup>55-11</sup> ts<sup>h</sup>ai<sup>213</sup>  
 1SG like 3SG cook REL dish  
 ‘I like the dishes he cooks.’ (Xu & Matthews, 2011, p. 114)
- b. hi<sup>33</sup> kai<sup>11</sup> [ua<sup>53</sup> t<sup>h</sup>aŋ<sup>11</sup> tsa<sup>53</sup> k<sup>h</sup>ɔi<sup>2</sup> li<sup>11</sup> tsia<sup>2</sup>] kai<sup>11</sup> bar  
 that CLF 1SG just.now give 2SG eat REL bar  
 ‘the bars I let you eat just now’ (Low, 2014, p. 27)

Conversely, Peng (2011) showed how Jambi Teochew Action Modification constructions can use the post-ObjR strategy with the use of a “relativizer”<sup>73</sup> *yang*, borrowed from Malay, as in (79).

- (79) nongkiã yang [Aling p<sup>h</sup>a? (kai)] k<sup>h</sup>au.  
 child REL Aling hit REL cry  
 ‘The child that Aling hit cried.’ (Peng, 2011, p. 262)

As first noted in McFarland (2017, p. 31), there is evidence that some speakers use the post-objR strategy with the linker strategy in CT Action Modification Constructions, as in (80).

- (80) a. 鸟 在 近 细 嫲孳团 个.. [食 汤]...  
 tɕiau<sup>52</sup> tɔ<sup>11</sup> kiŋ<sup>24</sup> soɪ<sup>21-53</sup> tɕa<sup>33</sup>nou<sup>33</sup>kiã<sup>52</sup> kai<sup>55-11</sup>.. [tɕia<sup>2</sup> tuŋ<sup>33</sup>]...  
 bird be.at near small woman-child REL.. eat soup..  
 ‘[The] bird is near the small girl who eats soup...’ [G2F1US]

---

<sup>73</sup> I don’t know enough about this word to say if it is a linker so we will use Peng’s term.

- b. 鸟 在 大夫团 个 [口 个 索]  
 tɕiau<sup>52</sup> tɔ<sup>11</sup> ta<sup>33</sup>pou<sup>33</sup>-kiã<sup>52</sup> kai<sup>55-11</sup> [nim<sup>24-11</sup> kai<sup>55-11</sup> sɔŋ<sup>2</sup>]  
 bird be.at man-child REL hold IK rope  
 ‘[The] bird is at [the] man who holds the rope.’ [G2F1US-PS]

The examples in (80) are from a session I had with G2F1US using picture stimuli from Lau (2016) designed to provoke different types of action modification constructions.

Unfortunately, I was not able to perform a similar procedure with the consultants in Cambodia as originally intended to determine the extent of this post-objR strategy. Future research would benefit from obtaining additional examples of CT Action Modification Constructions. The methodology and prompts from Lau (2016) are a potentially good way to procure such data.

The examples in (78-80) show action modification constructions with a shared participant to both events. Other Jieyang/Singapore Teochew Action Modifier constructions without a shared participant were also shown to use the pre-objR with linker strategies. This is the case for spatial or temporal relational head objRs, and place objRs, where the head objR does not corefer with an argument in the modifier clause. In (81a) from Jieyang and (81) from Singapore Teochew, 时候 *si<sup>11</sup>hau<sup>11</sup>* ‘time’ is the head ObjR, but it is not an argument of the modifier clause and the same is true for when *ti<sup>11</sup>hŋ<sup>33</sup>* ‘place’ is the modified ObjR as in (82) from Jieyang.

- (81) a. [ua<sup>53</sup> t’aŋ<sup>5-2</sup> toŋ<sup>33</sup> - haŋ<sup>5</sup>] kai<sup>55-11</sup> si<sup>53-35</sup>hou<sup>11</sup>  
 1SG study middle-school REL time  
 ‘(the time) when I was in high school’ (Xu, 2007, p. 115).

- b. [ʔ<sup>53</sup> naŋ<sup>11</sup>    tsɔ<sup>53</sup>    nou<sup>35</sup> kĩa<sup>53</sup>]    kai<sup>11</sup>    si<sup>11</sup> tsuŋ<sup>11</sup>    nɛ<sup>55</sup>  
 1PL.EXCL    be    children    MOD    time    PRT  
 kɛ<sup>33</sup> lai<sup>35</sup>    bɔ<sup>11</sup>    lui<sup>33</sup>    t<sup>h</sup>ak<sup>11</sup>    tsɿ<sup>33</sup>  
 home    NEG.have    money    read    book  
 ‘When we were children, we didn’t have money at home to study.’  
 [Low, 2014, p. 49]

- (82) [k’eʔ<sup>2</sup> naŋ<sup>55-11</sup>    paŋ<sup>213-53</sup>    k’a<sup>33</sup>ts’ia<sup>33</sup>    kai<sup>55-11</sup>    ti<sup>11</sup>hŋ<sup>33</sup>  
 give    people    put    bike    REL    place  
 ‘Places (or the place) where people can keep their bikes’ (Xu, 2007, p. 115).

At least for some speakers, CT Action Modification Constructions without a shared participant, such as with 时候 *si<sup>11</sup>hau<sup>11</sup>* ‘time’ as the head ObjR, also use the pre-objR with linker strategies, as in the examples in (83) from speaker G1F4.

- (83) a. ... [只个    孛团    死]    个    时候    □  
 ... [tɕia<sup>53</sup>    nou<sup>33</sup>kiã<sup>52</sup>    uk<sup>5</sup>]    kai<sup>55-11</sup>    si<sup>11</sup>hau<sup>11</sup>    nɔ<sup>55</sup>  
 ...    this.LW    child    sleep    REL    time    PRT  
 ‘...when this child sleeps.’ [G1F4-16]
- b. [伊    欲    死]    个    时候    伊    妈妈  
 [i<sup>33</sup>    ai<sup>21-53</sup>    uk<sup>5</sup>]    kai<sup>55-11</sup>    si<sup>11</sup>hau<sup>11</sup>    i<sup>33</sup>    ma<sup>55-11</sup>ma<sup>55</sup>  
 3SG    want    sleep    LW    time    3SG.POSS    mother  
 就    讲    故事    互    伊    听  
 tɕiu<sup>11</sup>    kaŋ<sup>53-24</sup>    ku<sup>53</sup>su<sup>11</sup>    hɔu<sup>11</sup>    i<sup>33</sup>    t<sup>h</sup>iã<sup>33</sup>  
 then    tell    story    give    3SG    listen  
 ‘When he wants to sleep, then his mother tells him a story and he listens.’ [G1F4-16]

There is some evidence that there is a reduction in the phonetic form of *kai<sup>55</sup>* when it acts as the linker in CT Action Modification Constructions, such as in (84) from G1F4 and in (85) from G1F1.

(84) [底畔 在 坐 拍字机] 个 时候 ...  
 [tɔi<sup>52-24</sup>paĩ<sup>55</sup> tɔ<sup>11</sup> tɕɔ<sup>24-11</sup> p<sup>h</sup>aɪ<sup>2-5</sup>dzi<sup>11</sup>ki<sup>33</sup>] ai<sup>11</sup> si<sup>11</sup>hau<sup>11</sup> ...  
 inside PROG sit type-machine LW time ...  
 ‘When [she] is sitting inside at the computer...’ [G1F4-16]

(85) a. 只个 [人 饲 猪] 个 地方 啊  
 tɕai<sup>55</sup> [naŋ<sup>55-11</sup> tɕ<sup>h</sup>i<sup>11</sup> tu<sup>33</sup>] a<sup>11</sup> ti<sup>11</sup>huŋ<sup>33</sup> a<sup>33</sup>  
 PDCO people raise pig LW place PRT

‘This (one) [is] a place where people raise pigs.’ [G1F1-8]

b. 好 只个 伊 坐 游泳池 口, [洒水] 个 地方 啊  
 hɔ<sup>52</sup> tɕi<sup>52</sup>kai<sup>55</sup> i<sup>33</sup> tɕɔ<sup>24-11</sup> iu<sup>11</sup>iɔŋ<sup>53</sup>ti<sup>55</sup>nia<sup>33</sup>, [siu<sup>55-11</sup>tɕui<sup>52</sup>] a<sup>11</sup> ti<sup>11</sup>huŋ<sup>33</sup> a<sup>33</sup>  
 good PDCO 3SG sit pool PRT, swim LW place PRT

‘Then [in] this (one [the picture]), he sits [at] a pool, the swimming place.’ [G1F1-8]

c. 在 耍 [[游]水] 个 地方 啊  
 tɔ<sup>11</sup> suŋ<sup>52</sup> [iu<sup>55-11</sup>]tɕui<sup>52</sup>] a<sup>11</sup> ti<sup>11</sup>huŋ<sup>33</sup> a<sup>33</sup>  
 PROG play swim REL place PRT

‘[People] are playing [at] the swimming place.’ [G1F1-8]

Additionally, there are other examples with 时候 *si<sup>11</sup>hau<sup>11</sup>* ‘time’ as the head ObjR that use the juxtaposition strategy. There is no linker word at all, as in (86).

(86) a. [伊 死] 时候 就 有 鹹酸 味  
 [i<sup>33</sup> ɿ<sup>24</sup>] si<sup>11</sup>hau<sup>11</sup> tɕiu<sup>11</sup> u<sup>11</sup> kiam<sup>11</sup>-suŋ<sup>33</sup> bi<sup>11</sup>  
 3SG sleep time then there is salt-sour smell

‘When he is sleeping, then there is the smell of food.’ [G3F1-3]

b. [伊 行 出] 时候 就 伊 想着.. 食 鱼  
 [i<sup>33</sup> kiã<sup>55-11</sup> tɕ<sup>h</sup>uk<sup>2</sup>]si<sup>11</sup>hau<sup>11</sup> tɕiu<sup>11</sup> i<sup>33</sup> siɔ<sup>24-11</sup> tiɔ<sup>25-2</sup>.. tɕiaɪ<sup>25-2</sup> hu<sup>55</sup>  
 3SG walk out time then 3SG think RC.. eat fish

‘When he left, then he thought of eating fish.’ [G3F1-3]

- c. 伊 听 [伊 在 死] 时候 啊  
 i<sup>33</sup> tʰiã<sup>33</sup> [i<sup>33</sup> tɔ<sup>11</sup> ʔ<sup>24</sup>] si<sup>11</sup> hau<sup>11</sup> a<sup>33</sup>  
 3SG listen 3SG PROG sleep time PRT  
 ‘He listens when he is sleeping.’ [G1F1-10]

Based on the evidence in the current data set, I would hypothesize that we are seeing a progression from (83) with the full linker *kaɪ*<sup>55</sup> to (84) and (85) with a linker with a reduced phonetic form, *ai*<sup>11</sup> or *a*<sup>11</sup>, to (86) where there is no linker present at all.<sup>74</sup> However, more data is needed to determine if there is indeed such a progression and to what extent the different strategies (linker, reduced linker, juxtaposition) are used by different speakers. It would also be beneficial to pay attention to the head ObjR in the construction to determine if the type of objR has any bearing on the strategy accepted or used.

A final type of CT Action Modification Construction shows lexical innovations or lexicalized CT Common ObjRs, similar to the CT Object Modification Constructions shown in §6.3.2. Examples of these using the compounding strategy are provided in (90).

---

<sup>74</sup> Croft (personal communication, May 30, 2022) noted that it appears here that the reduced forms of the linker only arise in what he would call CT “Noun Complement” Constructions. If it was the case that the reduced forms only appeared in one subtype of action modification constructions, this could justify differentiating between constructions with and without a shared participant, contrary to the approach taken here and in LaPolla (2017). It could also act as evidence that the strategy for CT “Noun Complement” Constructions was recruited for CT “Relative Clause” Constructions and the linker in the “Noun Complement” Constructions is being reduced because it has been in use longer in that construction. This is interesting because it is contrary to what is proposed in Croft (forthcoming, pp. 555-560). However, given the limited data I found on CT Action Modification Constructions with a shared participant, I am not confident saying that there are no examples of a linker with a reduced form in those constructions and will therefore not be making that argument here.

- |      |    |  |    |   |
|------|----|--|----|---|
| (90) | a. | 位泅水<br>ui <sup>11</sup> -siu <sup>55-11</sup> tɕui <sup>52</sup><br>place-swim<br>'pool' [G3F1-1]                                | b. | 位作田<br>ui <sup>11</sup> -tɕɔŋ <sup>5</sup> -tɕ <sup>h</sup> aŋ <sup>55</sup><br>place-make-field<br>'farm' [G3F1-1]                               |
|      | c. | 位煮碱酸<br>ui <sup>11</sup> -tɕu <sup>52-11</sup> -kiam <sup>11</sup> suŋ <sup>33</sup><br>place-cook-food<br>'kitchen' [G3F1-2, 3] | d. | 位洗衫裤<br>ui <sup>11</sup> -sɔi <sup>52-11</sup> -sã <sup>33</sup> k <sup>h</sup> ɔu <sup>21</sup><br>place-wash-clothes<br>'laundry room' [G3F1-2] |
|      | e. | 拍字机<br>p <sup>h</sup> aɪ <sup>2-5</sup> -dʒi <sup>11</sup> -ki <sup>33</sup><br>hit-letter-machine<br>'computer' [G1F4-16]       | f. | 飞机<br>pue <sup>33</sup> ki <sup>33</sup><br>fly-machine<br>'plane' [G3F1-1, G1F4-16]  |

## 6.5 Summary

This chapter covered the different ways object, properties, and actions can act as modifiers to referents in Cambodian Teochew. We have highlighted several things that seem to differentiate Cambodian Teochew from other Teochew varieties.

In §6.2.1, we saw evidence for the simplification of classifier system through the prevalence of the linker strategy for CT Enumeration Modification Constructions. It was also shown that a classifier or linker is not strictly required for CT Enumeration Modification Constructions or CT Deictic Attributive Constructions. Additionally, we presented evidence for a non-canonical word order in CT Enumeration Modification Constructions (ObjR + EC + (Classifier)) that mirrors the word order in Khmer. Cambodian Teochew may be experiencing a shift in acceptable word order for such constructions, which seems to provide evidence against Greenberg's (1990) hypothesis on shifting word order in enumeration modification constructions. In §6.2.2 we argued for the prevalence of post-ObjR modification with quantifier 'many'. In §6.3.1, we stated that we found no evidence for the use of the classifier or special form strategies for CT Possessive Constructions, though they are found in other varieties.

Finally, in §6.4, we showed the existence of some CT Action Modification Constructions using the post-ObjR strategy. There were also CT Action Modification Constructions that used a reduced form of the linker *kar*<sup>55</sup>, and some constructions that used the juxtaposition strategy, without the use of the linker *kar*<sup>55</sup> at all.

Overall many of the stated differences between Cambodian Teochew and other varieties show an increasing tendency for post-ObjR modification. Given that modifiers in Khmer are primarily ObjR-Modifier word order, it is hypothesized that contact with Khmer is causing these word orders changes in Cambodian Teochew.

The findings here should be considered preliminary and more research should be done to ascertain the prevalence of any of these patterns in the wider speaker community.

Additionally, there were many constructions we were not able to cover here due to lack of sufficient data on them. Some areas to focus for future work on include the use of admodifiers in property modification, the use of ordinal numbers, the existence and prevalence of mensural terms, possessive constructions with first and second person objects, particularly of kinship terms, and action modification constructions.

## Chapter 7 – Predication and Clauses

The information packaging function of predication expresses what the speaker is proclaiming about the referents in a certain utterance. The prototypical action predication asserts that a particular referent (or referents) is performing said action. CT Action Predication Constructions are covered in §7.1. Object predication states that the referent belongs to that object type. We cover CT Object Predication Constructions in §7.2. This section also covers two nonpredicational ways of packing information using object concepts. Property predication proclaims that the referent can be described by said property. CT Property Predication Constructions are explored in §7.3. In §7.4 we look at CT Locative Predication Constructions, an additional type of non-prototypical predication that asserts that a referent is at a particular location. A summary of the chapter is provided in §7.5.

### 7.1 CT Action Predication Constructions

The prototypical predication is action predication. Action predication constructions are made up of two types of items: the predicate, which is the head of the construction, and the arguments. The arguments of the predicate are referring phrases, which are used to encode the real-world referents. Semantically, these are the participants of the event. Croft differentiates between the term “argument” which describes the information packaging function, and “argument phrase” which is a comparative concept or construction that represents the combination of the participant and the argument (Croft, forthcoming, p. 156).

When linguists talk about argument phrases, they nearly always use terminology such as “subject” and “object”. Following LaPolla (1993, 2006), we will argue against the use of these

terms for Cambodian Teochew.<sup>75</sup> Evidence for this comes from the lack of a pivot in cross-clause coreference, which we will show below. First we need to discuss participant roles in events.

In regards to the number of participant roles in an event, there are traditionally thought to be three types of events: monovalent events with one core participant role, bivalent events with two core participant roles, and trivalent events with three core participant roles (Croft, forthcoming, p. 160). Prototypically, these three types of events lend themselves to three constructions: intransitive constructions with one core argument phrase, transitive constructions with two core argument phrases, and ditransitive constructions with three core argument phrases (Croft, forthcoming, pp. 160-1).

Some CT Intransitive Constructions are shown in (1), with the respective core argument underlined. Each of these has one core argument phrase. In (1a), 路 *lou<sup>11</sup>* ‘road’ is not a central participant role of the event of walking. In Chapter 5, we showed the frequent use of Zero Anaphora for active referents. Thus, the  $\emptyset$  in (1a) and (1c) should not be taken to mean that there is no core participant. Rather, the core participant is represented by  $\emptyset$ .

- (1) a. □ 猫 行 在 路.  $\emptyset$  行 在 路  
 u<sup>33</sup> ŋiau<sup>33</sup> kiã<sup>55-11</sup> tɔ<sup>11</sup> lou<sup>11</sup>.  $\emptyset$  kiã<sup>55-11</sup> tɔ<sup>11</sup> lou<sup>11</sup>  
 INTJ cat walk LOC road 3SG walk LOC road  
 ‘The cat walks on the road. [It] walks on the road.’ [G1F1-9]

---

<sup>75</sup> Croft (forthcoming, p. 158) used the terms “subject” and “object” to refer to the most and next most salient core arguments, respectively, in an action predication construction. LaPolla’s (1993, 2006) critique and mine are not directed at this different usage of said terms. However, similar to my argument against the use of traditional terms like “noun” and “pronoun”, I believe it can be confusing to the reader to use such traditional terminology.

- b. 猫            死  
n̩iau<sup>33</sup>        uk<sup>5</sup>  
cat            sleep  
‘A cat sleeps.’ [G1F2-14]
- c. ∅    跋    了  
∅    puaŋ<sup>5</sup>    ou<sup>11</sup>  
3SG    fall    already  
‘[The cat] has fallen.’ [G1F2-14]

Some CT Transitive Constructions are shown in (2), again with core arguments underlined. Each of these predicates has two core argument phrases. While Croft (forthcoming) encouraged grammar writers to show the use of the predicate ‘break’ as the prototypical transitive predicate, there were no tokens in the current data set. Instead, (2a) shows another change of state predicate (per Levin, 1993), 酸 *suŋ*<sup>33</sup> ‘cook’. The predicate in CT Transitive Constructions is expected to undergo tone sandhi as long as the second core argument is not a CT Personal Contextual ObjR Construction (Zhang, 1992, p. 202). This is seen to occur with the predicates 食 *tɕia*<sup>55-2</sup> ‘eat/drink’ and 拍 *p<sup>h</sup>a*<sup>2-5</sup> ‘hit’ in (2b). However, there were some cases in which the expected tone sandhi did not occur, such as with the predicate 抱 *p<sup>h</sup>ɔ*<sup>24</sup> ‘carry/hold’ in (2c).

- (2) a. 只个 伊            在    酸    也个  
tɕei<sup>55</sup>    i<sup>33</sup>                to<sup>11</sup>    suŋ<sup>33</sup>    mi<sup>5</sup>kai<sup>55</sup>  
PDCO    3SG                PROG    cook    something  
‘[In] this (one [the picture]), he is cooking something.’ [G1F1-10]

- b. 伊 食 水 □- 食 酒水  
*i*<sup>33</sup> *tɕiaŋ*<sup>5-2</sup> *tɕui*<sup>52</sup> *e*<sup>33</sup>- *tɕiaŋ*<sup>5-2</sup> *tɕiu*<sup>52</sup>*tɕui*<sup>52</sup>  
 3SG drink water INTJ drink alcohol  
好 伊 就 拍 伊 妘  
*hɔ*<sup>52</sup> *i*<sup>33</sup> *tɕiu*<sup>11</sup> *p<sup>h</sup>aŋ*<sup>2-5</sup> *i*<sup>33</sup> *bɔu*<sup>52</sup>  
 good 3SG then hit 3SG.POSS wife  
 ‘He drinks water- drinks alcohol, and then hits his wife.’ [G3F2-4]
- d. 伊 抱 伊个 团 □  
*i*<sup>33</sup> *p<sup>h</sup>ɔ*<sup>24</sup> *ia*<sup>31</sup> *kiã*<sup>52</sup> *ni*<sup>55</sup>  
 3SG carry 3SG.POSS child PRT  
 ‘She carries her child.’ [G1F2-13]

A prototypically monovalent event can become a bivalent event through the use of a causative construction. There was one example of a CT Causative Construction using a complex predicate strategy (Croft, forthcoming, p. 251), given in (3). We showed in (1c) that the predicate 跋 *puaŋ*<sup>5</sup> ‘fall’ is prototypically used for a monovalent event where one participant performs the action of ‘falling’. In (3), however, there are two participants, ‘the mother’ from the discourse context, represented by the first CT Zero Anaphora Construction  $\emptyset$ , and 猫物 *ŋiau*<sup>33</sup> *mueŋ*<sup>5</sup> ‘cat thing’. The causative relationship is expressed by the predicate 互 *hou*<sup>11</sup> which lexically means ‘give’.

- (3) ... $\emptyset$  来.. 拍 猫物 互  $\emptyset$  跋.. 顶- 跋 入 水  
 ... $\emptyset$  *lai*<sup>55-11</sup> .. *p<sup>h</sup>aŋ*<sup>2-5</sup> *ŋiau*<sup>33</sup> *mueŋ*<sup>5</sup> *hou*<sup>11</sup>  $\emptyset$  *puaŋ*<sup>5-2</sup> .. *teŋ*<sup>52</sup>- *puaŋ*<sup>5</sup> *dzip*<sup>5-2</sup> *tɕui*<sup>52</sup>  
 ...3SG come hit cat thing give 3SG fall.. atop- fall into water  
 ‘[The mother] comes and hits the cat thing, making [him] fall, fall into the water.’ [G3F1-3]

This use of a word that lexically means ‘give’ for causative contexts has been shown for Hokkien, Thai, Khmer, and others (Yap & Iwasaki, 2003, p. 424). However, according to Xu &

Matthews (2013, p. 240), an interpretation like this should not be possible in Chaozhou dialects.<sup>76</sup> They provided the example in (4) with *k<sup>h</sup>i?* ‘give’.

- (4)    *k<sup>h</sup>i?*    *i*        *pua?*    *lo?*    *k<sup>h</sup>u*  
          give    3SG    fall    down    go  
          ‘\*Make him fall down.’ (Xu & Matthews, 2013, p. 240)<sup>77</sup>

That the causative interpretation is allowed in Cambodian Teochew in (3), yet not in at least one other variety as in (4), could be showing an influence from Khmer, since that language allows these types of constructions. However, since there was only one token of this in the current data set, more data is needed to say anything more definite.

CT Ditransitive Constructions are shown in (5), with the predicate  $\overline{\text{互}}$  *hɔu<sup>11</sup>* ‘give’ taking three core argument phrases. The word order in (5a) and (5b) is agent + ‘give’ + recipient + item. (5c) shows agent + ‘give’ + item(s) + *k<sup>h</sup>u<sup>21</sup>* ‘go’ + recipient. It is unclear if the *k<sup>h</sup>u<sup>21</sup>* ‘go’ is some sort of special marking for the recipient in constructions with this order. In the examples in (6), one of the core arguments can be recovered from the context, so the order of the participant roles is unknown. Ultimately more data is needed to learn more about the strategies used for CT Ditransitive Constructions with  $\overline{\text{互}}$  *hɔu<sup>11</sup>* ‘give’.

- (5)    a.    嬉妘囡        物        就     $\overline{\text{互}}$     只个猫        碱酸        食  
          *tɕa<sup>33</sup>bɔu<sup>33</sup>kiã<sup>52</sup>* *mue?*<sup>5</sup>    *tɕiu<sup>11</sup>*    *hɔu<sup>11</sup>*    *tɕia<sup>53</sup>*    *ɲiau<sup>33</sup>*    *kiam<sup>11</sup>suŋ<sup>33</sup>*    *tɕia?*<sup>5</sup>  
          woman-child thing    then    give    this.LWcat    food        eat  
          ‘The girl thing then gives this cat food to eat.’ [G3F1-3]

<sup>76</sup> Xu and Matthews (2013) don’t limit their paper to a specific Teochew variety though Xu often writes on Jieyang.

<sup>77</sup> Transcription and glossing have been updated slightly.

- b. 伊 来 伊个 主 互 伊 饭 食 啊.  
 i<sup>33</sup> lai<sup>55</sup> ia<sup>31</sup> tɕu<sup>52</sup> həu<sup>11</sup> i<sup>33</sup> puŋ<sup>11</sup> tɕiaŋ<sup>5</sup> a<sup>55</sup>.  
 3SG come 3SG.POSS owner give 3SG rice eat PRT  
 Ø 互 伊 饭 食  
 Ø həu<sup>11</sup> i<sup>33</sup> puŋ<sup>11</sup> tɕiaŋ<sup>5</sup>  
 3SG give 3SG rice eat

‘He comes and his owner gives him rice to eat. [She] gives him rice to eat.’ [G1F2-14]

- c. 只个 马打 伊 互 口.. 只 个 鞋 𐄂𐄂 合 裤  
 tɕia<sup>53</sup> ma<sup>33</sup>ta<sup>24</sup> i<sup>33</sup> həu<sup>11</sup> a<sup>33</sup>.. tɕi<sup>52</sup> kai<sup>55-11</sup> ɕi<sup>55</sup> hei<sup>33</sup> kaŋ<sup>2-5</sup> k<sup>h</sup>ɕu<sup>21</sup>  
 this.LW police 3SG give INTJ.. this LW shoe and and pants  
 𐄂𐄂 合 衫 去 只 个 口.. 𐄂𐄂 or prisoner  
 hei<sup>33</sup> kaŋ<sup>2-5</sup> sã<sup>33</sup> k<sup>h</sup>u<sup>21-53</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> um.. neak tooh or prisoner  
 and and shirt go this LW INTJ.. prisoner or prisoner

‘This police officer, he gives these shoes and these pants and this shirt to this prisoner.’ [G3F2-4]

- (6) a. Ø 欲 主人 互 Ø 饭 食  
 Ø ai<sup>21</sup> tɕu<sup>52-24</sup> dzin<sup>55</sup> həu<sup>11</sup> Ø puŋ<sup>11</sup> tɕiaŋ<sup>5</sup>  
 3SG want master give 3SG rice eat  
 ‘[It] wants [its] master to give [it] rice to eat.’ [G1F1-9]
- b. 马打 就 互 Ø 伊 个 衫裤  
 ma<sup>33</sup>ta<sup>24</sup> tɕiu<sup>11</sup> həu<sup>11</sup> Ø i<sup>33</sup> kai<sup>55-11</sup> sã<sup>33</sup>k<sup>h</sup>ɕu<sup>21</sup>  
 police then give 3SG 3SG LW clothes  
 ‘The police officer then gives [him] his clothes.’ [G3F2-4]
- c. 只个 伊 挈 信 互 只, 阿老..老白人 Ø  
 tɕi<sup>52</sup>kai<sup>55</sup> i<sup>33</sup> k<sup>h</sup>ioŋ<sup>5-2</sup> sin<sup>21</sup> həu<sup>11</sup> tɕei<sup>55</sup>, a<sup>33</sup>-lau<sup>24-11</sup>..lau<sup>24-11</sup>-peŋ<sup>5</sup>-nan<sup>55</sup> Ø  
 PDCO 3SG take letter give PDCO NVOC-old..old-white-person 3SG  
 ‘[In] this (one [the picture]), he takes a letter and gives [it] to this one [pointing], an old person.’ [G1F1-8]

Similar to how a monovalent event could become bivalent with the use of 互 *hou*<sup>11</sup>

‘give’, a bivalent event can be construed as a trivalent event using 互 *hou*<sup>11</sup> ‘give’ to introduce a beneficiary, as in (7).

- (7) a.  $\emptyset$  工 去.. 挈 鐮.. 挈 鐮..  
 $\emptyset$  k<sup>h</sup>aŋ<sup>33</sup> k<sup>h</sup>u<sup>21-53</sup>.. k<sup>h</sup>iəʔ<sup>5-2</sup> lui<sup>33</sup>.. k<sup>h</sup>iəʔ<sup>5-2</sup> lui<sup>33</sup>..  
 3PL work go take money take money  
 互 团.. 去 读书  
 hou<sup>11</sup> kiã<sup>52</sup>.. k<sup>h</sup>u<sup>21-53</sup> t<sup>h</sup>ak<sup>2</sup>tɕu<sup>33</sup>  
 give child go school  
 ‘[They] go to work to make money, make money for [their] child to go to school.’ [G3F2-4]
- b. 伊 欲 死 个 时候 伊 妈妈  
 i<sup>33</sup> ai<sup>21-53</sup> uk<sup>5</sup> kai<sup>55-11</sup> si<sup>11</sup>hau<sup>11</sup> i<sup>33</sup> ma<sup>55-11</sup>ma<sup>55</sup>  
 3SG want sleep REL time 3SG.POSS mother  
 就 讲 故事 互 伊 听  
 tɕiu<sup>11</sup> kaŋ<sup>53-24</sup> ku<sup>53</sup>su<sup>11</sup> hou<sup>11</sup> i<sup>33</sup> t<sup>h</sup>iã<sup>33</sup>  
 then tell story give 3SG listen  
 ‘When he wants to sleep, then his mother tells him a story and he listens.’ [G1F4-16]
- c. 只 个 老 伊 在 但 互 只 个 丈夫  
 tɕi<sup>52</sup> kai<sup>55-11</sup> lau<sup>24</sup> i<sup>33</sup> tɔ<sup>11</sup> tã<sup>21-53</sup> hou<sup>11</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> ta<sup>33</sup>pou<sup>33</sup>  
 this LW old 3SG PROG tell give this LW man  
 ‘This old man is telling this man.’ [G3F2-4]

Now we can return to the issue of “subject” and “object” terminology. Some languages show an alignment strategy, or syntactic pivot, in that different arguments of intransitive and transitive constructions behave similarly. The three relevant arguments are: S, the single participant role of intransitive predicate; A, the prototypical agent or agent-like participant role of transitive predicate; and P, the patient or patient-like participant of transitive predicate. In the accusative alignment strategy, S and A pattern together, while in the ergative alignment

strategy, S and P pattern together. LaPolla (1993) provided evidence from cross-clause coreference, relativization, *bi* comparatives, raising to subject, indispensability, reflexives, and pseudo-passives to show that there was no such alignment strategies or syntactic pivot in Mandarin. Thus, there is no evidence that would support an argument for the concept of “subject” in Mandarin. Evidence from cross-clause reference suggests that these same arguments can be made for Cambodian Teochew.<sup>78</sup>

In English, a language with an accusative alignment strategy, or [S, A] pivot, when two clauses are coordinated, the English Zero Anaphora Construction,  $\emptyset$ , can only be used to refer to a referent that is in the S or A role in both clauses, as in (8). (8c) is ungrammatical because English does not allow the S role of the first clause to corefer with the P role, the  $\emptyset$ , in the second clause.

- (8) a. The child stood and  $\emptyset$  saw the rooster.  
 b. The rooster stood and  $\emptyset$  was seen by the child.  
 c. \*The rooster stood and the child saw  $\emptyset$ .

In languages with an ergative alignment strategy, or [S, P] pivot, for coordination, such as Dyirbal,<sup>79</sup> a sentence like (8a) would not be allowed because the S and the A do not pattern together, while sentences like (8b) and (8c) are allowed.

In Cambodian Teochew, however, all of the equivalent sentences in (8) are allowed. The shared argument of a conjoined structure can appear as a CT Zero Anaphora Construction no matter if it is the A or the P role in the second clause. In (9), 鸡翁 *kɔŋ<sup>33</sup> aŋ<sup>11</sup>* ‘rooster’ is

---

<sup>78</sup> Future work could explore some of the other constructions covered in LaPolla (1993) in Cambodian Teochew to strengthen the argument.

<sup>79</sup> See examples in LaPolla (1993).

acting as the S of the predicate 企 *k<sup>h</sup>ia<sup>24</sup>* ‘stand’ in the first clause, and fulfills the P role of the predicate 睇 *t<sup>h</sup>ɔ̃<sup>52</sup>* ‘see’ in the second clause. Meanwhile, in (10), 伊 *i<sup>33</sup>* ‘3SG’ is in the S role for the predicate 行 *kiã<sup>55</sup>* ‘walk’, and is in the A role of the predicate 睇见 *t<sup>h</sup>ɔ̃<sup>52-11</sup>kiã<sup>21</sup>* ‘see’ in the second clause. Note that the second clause in both of these examples is not any special construction. They are CT Basic Voice Constructions.

- (9) 只 隻 鸡翁 企 在 篱笆 顶  
*tɕi<sup>52</sup> tɕiaŋ<sup>2-5</sup> kɔi<sup>33</sup>aŋ<sup>11</sup> k<sup>h</sup>ia<sup>24</sup> tɔ<sup>11</sup> li<sup>11</sup>pa<sup>33</sup> teŋ<sup>52</sup>*  
 this CLF.ANIMAL rooster stand LOC fence atop  
 孛团 在 企 睇 Ø  
*nɔu<sup>33</sup>kiã<sup>52</sup> tɔ<sup>11</sup> k<sup>h</sup>ia<sup>24-11</sup> t<sup>h</sup>ɔ̃<sup>52</sup> Ø*  
 kid PROG stand see 3SG

‘This rooster is standing on top of a fence and [some] kids are standing looking [at it].’ [G1F3-15]

- (10) 伊 口 行 就 Ø 睇见 獠\*鼠  
*i<sup>33</sup> na<sup>11</sup> kiã<sup>55</sup> tɕiu<sup>11</sup> Ø t<sup>h</sup>ɔ̃<sup>52-11</sup>kiã<sup>21-53</sup> ŋiau<sup>11</sup>tɕ<sup>h</sup>u<sup>52</sup>*  
 3SG PROG walk then 3SG see mouse  
 口 土墙 顶畔  
*na<sup>11</sup> t<sup>h</sup>ɔu<sup>11</sup>tɕ<sup>h</sup>iɔ<sup>55</sup> teŋ<sup>52-11</sup>paĩ<sup>55</sup>*  
 LOC wall atop

‘He is walking then [he] sees a mouse on top of a wall.’ [G3F1-3]

Given what we have just presented on Cambodian Teochew, we won’t talk about “subject”, or “direct object” or any type of alignment strategies since they are not viable concepts for this language.

The range of possibilities for the types of action predication constructions is immense. In this section, we try to highlight some of the most important/common CT Action Predication Constructions. The examples we have presented so far are all in the declarative, in that they are

asserting some propositional content (Croft, forthcoming, p. 331). In §7.1.1, we show the ways in which CT Declarative Constructions are negated. Two types of non-declarative constructions explored are CT Interrogative Constructions (§7.1.2) and CT Imperative-Hortative (§7.1.3). Finally, we cover CT Complex Predicate Constructions (§7.1.4) and complex sentences (§7.1.5).

### 7.1.1 CT Declarative Negation Constructions

This section focuses on the declarative negation construction (Croft, forthcoming, p. 337). Teochew languages have historically exhibited a wide number of morphemes that can be used in Teochew Declarative Negation Constructions. Some of these morphemes can be used for negation in other constructions, such as object or property predication. Their usages in those constructions are covered in the respective sections. The morphemes used for negation in declarative sentences are 唔  $m^{11}$  ‘NEG’ (§7.1.1.1), 无  $b\sigma^{55}$  ‘NEG’ (§7.1.1.2), 未  $bue^{11}$  ‘not.yet’ (§7.1.1.3), and 𠵼  $b\sigma^{24}$  ‘NEG.can’ (§7.1.1.4).

Overall, there were not an abundance of CT Declarative Negation Constructions in the collected data so it is hard to draw any firm conclusions on the usage of the different negation morphemes, especially since there are so many. Future research should focus on gathering CT Declarative Negation Constructions with a wide variety of predicates in order to identify and confirm patterns in their usages.

#### 7.1.1.1 唔 $m^{11}$ ‘NEG’

唔  $m^{11}$  ‘NEG’ is used in CT Declarative Negation Constructions with predicates of cognition such as 记/记得  $k\dot{r}^{21}/k\dot{r}^{21-53}tik^2$  ‘remember’, in (11).

- (11) a. 所以 恬 无 唔 记得  
 so<sup>24</sup>t<sup>52</sup> tiam<sup>11</sup> bə<sup>55-11</sup> m<sup>11</sup> ki<sup>21-53</sup>tik<sup>2</sup>  
 so always NEG NEG remember  
 ‘So [I] never forget.’ [G1F6-24]
- b. 伊 唔 记 路 了  
 i<sup>33</sup> m<sup>11</sup> ki<sup>21-53</sup> lou<sup>11</sup> ou<sup>11</sup>  
 3SG NEG remember road already  
 ‘He forgot the way.’ [G2F1US-18]

唔 *m<sup>11</sup>* ‘NEG’ is also typically used in CT Declarative Negation Constructions with another predicate of cognition, 知 *tcai<sup>33</sup>* ‘know’, as in (12).

- (12) a. 只个 唔 知 啲 来 X  
 tɕi<sup>52</sup>kai<sup>55</sup> m<sup>11</sup> tcai<sup>33</sup> tiaŋ<sup>55</sup> lai<sup>55</sup> nia<sup>33</sup>  
 PDCO NEG know who come PRT  
 ‘[In] this (one [the picture]), [I] don’t know who comes.’ [G2M1-6]
- b. 我 唔 知 唔是 意思  
 ua<sup>52</sup> m<sup>11</sup> tcai<sup>33</sup> mi<sup>24</sup> i<sup>53</sup>su<sup>33</sup>  
 1SG NEG know NEG.COP meaning  
 ‘I don’t know [what] is the meaning.’ [G1M1-7]
- c. 唔 知 团 七个 跔跔 啊是 七个 口 口  
m<sup>11</sup> tcai<sup>33</sup> kiã<sup>52</sup> mi<sup>25</sup>kai<sup>55</sup> k<sup>h</sup>u<sup>55</sup>k<sup>h</sup>u<sup>55</sup>k<sup>h</sup>u<sup>55</sup> a<sup>11</sup>si<sup>24</sup> mi<sup>25</sup>kai<sup>55</sup> nia<sup>11</sup> nia<sup>11</sup>  
NEG know child what crouch RED RED or what PRT RED  
 ‘[I] don’t know, the child, what, he’s crouching, or what.’ [G1F1-11]

There are several tokens from speakers G1F1, G1M1, G2M1, and G2F1US using 唔 *m<sup>11</sup>* ‘NEG’ in the CT Declarative Negation Construction with predicate 知 *tcai<sup>33</sup>* ‘know’. This same usage was noted in Singapore Teochew by Yeo (2011) and Low (2014, p. 39), in Jieyang Teochew (Xu, 2005, 2007), and in Pontianak Teochew (Veniranda, 2015, p. 105). However, there is evidence that some CT speakers say 无知 *bə<sup>11</sup>tcai<sup>33</sup>* ‘not know’ as in the series of

utterances from G3F2 and G3F3 in (13). More evidence is needed to see the extent of this usage.

- (13) a. 先回 伊 怎呢 无.. 无 知.. 无 知 但—  
 sai<sup>33</sup>hue<sup>55</sup> i<sup>33</sup> tɕɔ<sup>52</sup>ni<sup>55</sup> bɔ<sup>11</sup>.. bɔ<sup>11</sup> tɕai<sup>33</sup>.. bɔ<sup>11</sup> tɕai<sup>33</sup> tã<sup>21-53</sup>—  
 before 3SG why NEG.. NEG know.. NEG know COMP—  
 ‘Why didn’t he know before that...’ [G3F3-4]
- b. 无 知 好好  
 bɔ<sup>11</sup> tɕai<sup>33</sup> hɔ<sup>52</sup>hɔ<sup>52</sup>  
 NEG know good RED  
 ‘Not know [how good things were before].’ [G3F2-4]
- c. 无 知 好  
 bɔ<sup>11</sup> tɕai<sup>33</sup> hɔ<sup>52</sup>  
 NEG know good  
 ‘Not know it was good.’ [G3F3-4]

Zhang (1996, p. 91) stated that in Chaoshan 晓 *hiau*<sup>52</sup> ‘know.how; able’<sup>80</sup> should be negated with 唔 *m*<sup>11</sup> ‘NEG’ and not 𠵹 *bɔi*<sup>24</sup> ‘NEG.can’. However, there are examples of both combinations being used in Cambodian Teochew, as in (14). There also seem to be examples of the same predicate negated with *bɔi*<sup>11</sup> in Singapore Teochew, as in (15). Veniranda (2015, p. 99) stated that this predicate is negated with *boi* in Pontianak Teochew. More examples of CT Declarative Negation Constructions with this predicate are needed.

- (14) a. 我 唔 晓 但 只个  
 ua<sup>52</sup> m<sup>11</sup> hiau<sup>53-24</sup> tã<sup>21</sup> tɕi<sup>52</sup>kai<sup>55</sup>  
 1SG NEG know.how say PDCO  
 ‘I don’t know how to say this.’ [G2M1-5]

---

<sup>80</sup> Zhang’s (1996) paper had no transcriptions or English meanings so I have added them.

- b. “汝 摘 伊 伊 就 歪 晓 笑”  
 “lu<sup>52</sup> tia<sup>2-5</sup> i<sup>33</sup> i<sup>33</sup> tɕiu<sup>11</sup> bɔi<sup>24-11</sup> hiau<sup>52-24</sup> tɕ<sup>hiɔ</sup><sup>21</sup>”  
 2SG pick 3SG 3SG then NEG.can able smile  
 “If you pick it, then it won’t be able to smile.” [G2F1US-19]

- (15) a. bɔi<sup>11</sup> hiou<sup>33</sup> t<sup>hiou</sup><sup>53</sup> tsui<sup>11</sup> kai<sup>11</sup> naŋ<sup>55</sup>  
 NEG.can jump water MOD person  
 ‘people who don’t know how to dive’ (Low, 2014, p. 27)
- b. ... i<sup>33</sup>naŋ<sup>11</sup> bɔi<sup>11</sup>-hiau<sup>53</sup> hiaɿ<sup>5</sup> k<sup>h</sup>a<sup>33</sup>ts<sup>h</sup>ia<sup>33</sup> kai<sup>11</sup>  
 ... 3PL NEG-know stop bicycle NMLZ  
 ‘...they don’t know how to park the bicycle.’ (Yeo, 2011, p. 98)

Similarly, Xu (2007, p. 240) noted that 唔 *m*<sup>35</sup> ‘NEG’ is used in Jieyang to negate the predicate 好 *hāu*<sup>213</sup> ‘be willing’, as in (16). There was only one token of this predicate in the collected data on Cambodian Teochew and it was negated with 无 *bɔ*<sup>55</sup> ‘NEG’, as in (17).

- (16) i<sup>33</sup> m<sup>35-21</sup> hāu<sup>213-53</sup> sio<sup>33</sup>-hu<sup>35</sup> ua<sup>53-213</sup>  
 3SG not willing help 1SG  
 ‘She’s not willing to (won’t) help me.’ (Xu, 2007, p. 240)

- (17) 咀 伊 伊 无 好 □  
 tã<sup>21-53</sup> i<sup>33</sup>, i<sup>33</sup> bɔ<sup>55-11</sup> hāu<sup>21</sup> nia<sup>33</sup>  
 say 3SG, 3SG NEG willing PRT  
 ‘He says he is not willing [to drink alcohol].’ [G3F2-4]

Xu (2005, p. 184) also noted that 唔 *m*<sup>35</sup> ‘NEG’ is used in Jieyang to negate the predicate 惊 *kiā*<sup>33</sup> ‘be afraid’, as in (18). However, speaker G2F1US negates 惊 *kiā*<sup>33</sup> ‘fear’ with 无 *bɔ*<sup>55</sup> ‘NEG’, shown in (19).

- (18) ua<sup>53</sup> m<sup>35-21</sup> kiā<sup>33</sup> tsi<sup>53-35</sup> kai<sup>55-11</sup> naŋ<sup>55</sup>  
 1SG not be.afraid this CLF person  
 ‘I am not afraid of this person.’ (Xu, 2005, p. 184)

- (19) a. “咋呢 羊.. 伊 无 惊 我”  
 “tɕɔ<sup>52</sup>ni<sup>55</sup> i<sup>55</sup>.. i<sup>33</sup> bɔ<sup>55-11</sup> kiã<sup>33</sup> ua<sup>52</sup>”  
 why sheep.. 3SG NEG fear 1SG  
 “‘Why does the sheep not fear me?’” [G2F1US-18]
- b. 伊 无 惊 只个 X. 伊 惊 老虎  
 i<sup>33</sup> bɔ<sup>55-11</sup> kiã<sup>33</sup> tɕi<sup>52</sup>kai<sup>55</sup> nia<sup>33</sup>. i<sup>33</sup> kiã<sup>33</sup> lau<sup>24-11</sup>hou<sup>52</sup>  
 3PL NEG fear PDCO PRT. 3PL fear tiger  
 ‘They don’t fear this (one). They fear the tiger.’ [G2F1US-18]

Finally, Xu (2005, p. 185, 2007, p. 238) and Zhuang (2001, p. 48) stated that *m*<sup>35</sup> can additionally be used to negate volition in Jieyang/Chaoshan, as in (20).

- (20) i<sup>33</sup> m<sup>35-21</sup> k<sup>h</sup>u<sup>213-53</sup> ha<sup>ʔ5-2</sup>hau<sup>35</sup>  
 3SG not go school  
 ‘He doesn’t/didn’t want to go to school.’ (Xu, 2005, p. 185)

There was no evidence for this usage in Cambodian Teochew. Low (2014, p. 38) also noted that *m*<sup>11</sup> is not used to negate volition in Singapore Teochew.

Based on the behavior of 唔 *m*<sup>11</sup> ‘NEG’ shown in this section, initial evidence points to the reduction in the usages of 唔 *m*<sup>11</sup> ‘NEG’ in CT Declarative Negation Constructions, sometimes in favor of 无 *bɔ*<sup>55</sup> ‘NEG’, or 𠵼 *bɔ*<sup>24</sup> ‘NEG.can’, though more data is needed to confirm/strengthen this hypothesis.

#### 7.1.1.2 无 *bɔ*<sup>55</sup> ‘NEG’

The morpheme 无 *bɔ*<sup>55</sup> ‘NEG’ is used in CT Declarative Negation Constructions to indicate the non-occurrence of event/activity, as in (21).

- (21) a. Ø 行 入.. 客厅, Ø 无 睇见  
 Ø kia<sup>55-11</sup> jip<sup>5-2</sup>.. k<sup>h</sup>eŋ<sup>5</sup>-t<sup>h</sup>uŋ<sup>55</sup>, Ø bɔ<sup>11</sup> t<sup>h</sup>ɔŋ<sup>52</sup>kiaŋ<sup>21</sup>  
 3SG walk into.. guest-room, 2SG NEG see  
 ‘[She] walked into the living room and didn’t see [him].’ [G3F1-2]
- b. 伊 水 觅 掠 鱼 掠 无 着  
 i<sup>33</sup> tɕui<sup>52</sup> tɕ<sup>h</sup>ue<sup>11</sup> liaŋ<sup>5-2</sup> hu<sup>55</sup> liaŋ<sup>5-2</sup> bɔ<sup>11</sup> tiɔŋ<sup>5</sup>  
 3SG water seek catch fish catch NEG RC  
 ‘He looked to catch fish in the water but didn’t catch any.’ [G3F1-3]
- c. 伊 掠 无 着 獠\*鼠 了  
 i<sup>33</sup> liaŋ<sup>5-2</sup> bɔ<sup>55-11</sup> tiɔŋ<sup>2</sup> ŋia<sup>24</sup>tɕ<sup>h</sup>u<sup>52</sup> ɔu<sup>11</sup>  
 3SG catch NEG RC mouse already  
 ‘It didn’t catch the mouse.’ [G1F1-9]
- d. 伊 无 睇 着  
 i<sup>33</sup> bɔ<sup>11</sup> t<sup>h</sup>ɔŋ<sup>52-24</sup> tiɔŋ<sup>2</sup>  
 3SG NEG see RC  
 ‘He didn’t see [the frog].’ [G1F1-11]
- e. Ø 无 咁 话 □  
 Ø bɔ<sup>55-11</sup> tã<sup>21-53</sup> ue<sup>11</sup> nia<sup>11</sup>  
 3SG NEG say word PRT  
 ‘[The flower] didn’t say anything.’ [G2F1US-19]

There was one usage 无 bɔ<sup>55</sup> ‘NEG’ indicating the negation of a habitual event, as in (22). Similar usage was noted in Jieyang by Xu (2007, p. 226) and in Singapore Teochew by Low (2014, p. 40). More data is needed from Cambodian Teochew to determine more about this usage.

- (22) 伊 咁 伊 无 食 □  
 i<sup>33</sup> tã<sup>21-53</sup> i<sup>33</sup> bɔ<sup>55-11</sup> tɕiaŋ<sup>5</sup> nia<sup>33</sup>  
 3SG say 3SG NEG eat PRT  
 ‘He says he doesn’t drink [alcohol].’ [G3F2-4]

Other Teochew varieties (Jieyang (Xu, 2007, p. 238), Johor Bahru (X. Chen, 2003, p. 387), Pontianak (Peng, 2012, p. 153), Singapore (Low, 2014, p. 41)) have noted the existence of a negative morpheme 𠵿 *maĩ*<sup>213</sup> ‘NEG.want’ (a contraction of 唔 *m*<sup>11</sup> + 爱 *aĩ*<sup>213</sup>) used in constructions denoting volition. While anecdotally I can speak to my consultants/family members using this morpheme, it was not present in the recorded data.

There is some evidence from G2F1US that at least some speakers use 无欲 *bɔ*<sup>55-11</sup>*aĩ*<sup>21</sup> ‘NEG want’, as in (23). Xu (2007, pp. 238-240) showed slightly different interpretations for utterances with 𠵿 *maĩ*<sup>213</sup> ‘NEG.want’ versus 无欲 *bɔ*<sup>55-11</sup>*aĩ*<sup>21</sup> ‘NEG want’ in Jieyang Declarative Negation Constructions. Additional research needs to be done to determine the functions of these morphemes in Cambodian Teochew. It would be expected that the difference between them would be roughly the same as the difference between the 唔 *m*<sup>11</sup> ‘NEG’ and 无 *bɔ*<sup>55</sup> ‘NEG’ generally (R.J. LaPolla, personal communication, December 27, 2021).

- (23) a. 伊 但 “我 无 欲 摘 汝 □”  
           *i*<sup>33</sup>   *tã*<sup>21</sup>   “*ua*<sup>52</sup> *bɔ*<sup>55-11</sup> *aĩ*<sup>21-53</sup> *tiaŋ*<sup>2-5</sup> *lu*<sup>52</sup> *nia*<sup>11</sup>”  
           3SG   say   1SG   NEG   want   pick   2SG   PRT  
           ‘He says, “I don’t want to pick you.”’ [G2F1US-19]
- b. 我 无 欲 食 汝 了  
           *ua*<sup>52</sup>   *bɔ*<sup>55-11</sup> *aĩ*<sup>21-53</sup> *tɕia*<sup>2-2</sup> *lu*<sup>52</sup>   *ɔu*<sup>11</sup>  
           1SG   NEG   want   eat   2SG   already  
           ‘I don’t want to eat you anymore.’ [G2F1US-18]

### 7.1.1.3 未 *bue*<sup>11</sup> ‘not yet’

CT Declarative Negation Constructions with 未 *bue*<sup>11</sup> ‘not.yet’ indicate that an event has not yet happened, when there is the expectation that it will happen in the future. In this regard, 未 *bue*<sup>11</sup> ‘not.yet’ expresses negation as well as modality. There were only two predicates 未 *bue*<sup>11</sup> ‘not.yet’ was found with, 醒 *tɕ*<sup>h</sup>*e*<sup>52</sup> ‘wake.up’, as in (24), and 来 *lai*<sup>55</sup> ‘come’, in (25).

There is an expectation that these events will come to fruition: 孛囡 *nou<sup>33</sup>kiã<sup>52</sup>* ‘child’ will wake up and 媛 *ai<sup>55</sup>* ‘mother’ will come. The pair of utterances in (25) show that 未 *bue<sup>11</sup>* ‘not.yet’ can appear before or after the predicate.

- (24) a. 只个 孛囡 伊 未 醒  
*tɕei<sup>55</sup> nou<sup>33</sup>kiã<sup>52</sup> i<sup>33</sup> bue<sup>11</sup> tɕ<sup>h</sup>ẽ<sup>52</sup>*  
 PDCO child 3SG not.yet wake.up  
 ‘[In] this (one [the picture]), the child, he hasn’t woken up.’ [G1F1-10]
- b. 孛囡 死. Ø 未 醒  
*nou<sup>33</sup>kiã<sup>52</sup> ʔ<sup>24</sup>. Ø bue<sup>11</sup> tɕ<sup>h</sup>ẽ<sup>52</sup>*  
 child sleep. 3SG not.yet wake.up  
 ‘[The] child sleeps. He hasn’t woken up yet.’ [G1F1-10]
- c. 个 孛囡 好 死. Ø 未 早醒  
*kai<sup>55-11</sup> nou<sup>11</sup>kiã<sup>52</sup> ho<sup>52-24</sup> uk<sup>5</sup>. Ø bue<sup>11</sup> tɕau<sup>52-24</sup>tɕ<sup>h</sup>ẽ<sup>52</sup>*  
 IK child good sleep. 3SG not.yet wake.up  
 ‘The child sleeps well. [He] hasn’t woken up.’ [G1F4-16]
- (25) a. 伊 睇 伊 媛 来 未  
*i<sup>33</sup> t<sup>h</sup>ɔ̃<sup>52</sup> i<sup>33</sup> ai<sup>55</sup> lai<sup>55</sup> bue<sup>11</sup>*  
 3SG see 3SG.POSS mother come not yet  
 ‘He looks to see if his mother has come or not.’ [G2M1-6]
- b. 好 伊 等 伊 媛 来.. Ø 未 来  
*ho<sup>52</sup> i<sup>33</sup> tan<sup>52-11</sup> i<sup>33</sup> ai<sup>55</sup> lai<sup>55</sup>.. Ø bue<sup>11</sup> lai<sup>55</sup>*  
 good 3SG wait 3SG.POSS mother come.. 3SG not yet come  
 ‘Then he waits for his mother to come... She hasn’t come yet.’  
 [G2M1-6]

Similar usages are noted in Singapore and Jieyang varieties in Low (2014, p. 41) and Xu (2007, p. 255).<sup>81</sup>

<sup>81</sup> See also Zhang (1999).

7.1.1.4 𠵹 *bɔi*<sup>24</sup> ‘NEG.can’

The use of 𠵹 *bɔi*<sup>24</sup> ‘NEG.can’ in a CT Declarative Negation Construction shows that the agent lacks the ability to perform the action, shown in (26). 𠵹 *bɔi*<sup>24</sup> ‘NEG.can’ is another example of a form that expresses negation as well as modality.

- (26) a. 伊 𠵹.. 伊 𠵹, 𠵹 carry 只 个 壳 啊  
*i*<sup>33</sup> *bɔi*<sup>24-11</sup>.. *i*<sup>33</sup> *bɔi*<sup>24-11</sup>, *bɔi*<sup>24-11</sup> carry *tɕi*<sup>52</sup> *kai*<sup>55-11</sup> *kʰak*<sup>2</sup> *a*<sup>33</sup>  
 3SG NEG.can.. 3SG NEG.can, NEG.can carry this LW shell PRT  
 ‘He cannot..he cannot carry this shell.’ [G2F1US-17]
- b. “怎呢 个 涂 合 天 𠵹.. 𠵹 相辅 俺?”  
*tɕɔ*<sup>52</sup>*ni*<sup>55</sup> *kai*<sup>55-11</sup> *tʰɔu*<sup>55</sup> *kaʔ*<sup>5</sup> *tʰɿ*<sup>33</sup> *bɔi*<sup>24-11</sup>.. *bɔi*<sup>24-11</sup> *siɔ*<sup>33</sup>*hu*<sup>24-11</sup> *naŋ*<sup>52</sup>”  
 why IK dirt and sky NEG.can.. NEG.can help 1PL  
 “‘Why can’t the dirt and sky help us?’” [G2F1US-17]
- c. “∅ 跳跳跳 𠵹 过”  
 “∅ *tʰiau*<sup>21-53</sup>*tʰiau*<sup>21-53</sup>*tʰiau*<sup>21</sup> *bɔi*<sup>24-11</sup> *kue*<sup>21</sup>”  
 1PL jump RED RED NEG.can cross  
 “‘[We] cannot jump across.’” [G2F1US-20]

Other usages of 𠵹 *bɔi*<sup>35</sup> ‘NEG.can’ in negating predicates were described in Xu (2005, 2007) and Zhang (1996) but they were not found in the current data set on Cambodian Teochew.

7.1.2 CT Interrogative Constructions

In interrogative constructions, some propositional content is unknown to the speaker. Croft (forthcoming, p. 341) outlined three types of interrogative constructions.

In polarity questions, the speaker is requesting the polarity (yes or no) of the proposition. CT Polarity Question Constructions can be formed with the negative word 无 *bɔ*<sup>55</sup> ‘NEG’ and the rising intonation strategy, as in (27).

- (27) a. 着 无? 着 无?  
 tioʔ<sup>5</sup> bɔ<sup>55</sup>? tioʔ<sup>5</sup> bɔ<sup>55</sup>?  
 right NEG right NEG  
 ‘Right? Right?’ [G1F2-13]
- b. 伊 但 “有 空 也个 无?”  
 i<sup>33</sup> tã<sup>21</sup> “u<sup>11</sup> k<sup>h</sup>aj<sup>33</sup> miʔ<sup>5</sup>kai<sup>55</sup> bɔ<sup>55</sup>?”  
 3SG say have hole what NEG  
 ‘He says, “Is there a hole or something?”’ [G2F1US-20]
- c. “俺 去 无”  
 “naŋ<sup>52</sup> k<sup>h</sup>u<sup>21</sup> bɔ<sup>55</sup>  
 1PL go NEG  
 “‘Let’s go, no?’” [G2F1US-20]
- c. Ari 放屎 无  
 Ari paŋ<sup>21-53</sup>sai<sup>52</sup> bɔ<sup>55</sup>  
 Ari poop NEG  
 ‘Did Ari Poop?’ [G2M2US]

Low (2014, p. 43) showed that other types of negative words can be used to form Singapore Polarity Question Constructions. No examples were found in the current study of other negative words used in CT Polarity Question Constructions, though there were not many tokens of this type of construction, so more data is needed in this regard.

No negative word is needed to form a CT Polarity Question Construction. There are some instances where the rising intonation strategy is enough to indicate that the construction is an interrogative, as the pair of utterances in example (28) shows.

- (28) a. 好  
 hó<sup>52</sup>  
 good  
 ‘Good?’ [G1M1-7]

- b.     好  
           hǎo<sup>52</sup>  
           good  
           ‘Good.’ [G1M1-7]

Another strategy used for polarity questions is the A-not-A strategy (Croft, forthcoming, p. 343). This strategy is commonly used in Sinitic languages. An example of a Jieyang Polarity Question Construction using the A-not-A strategy is given in (29). No examples of constructions with this strategy were found in the current data set so it remains a question for future research.

- (29)   lu<sup>53</sup>    si<sup>35-21</sup>-m<sup>35-21</sup>-si<sup>35-21</sup>    lau<sup>55-11</sup>su<sup>33</sup>  
           2SG    COP-not-COP        teacher  
           ‘Are you a/the teacher?’ (Xu, 2007, p. 254)

The second type of interrogative construction is information questions (Croft, forthcoming, p. 341). The open proposition in these constructions is some piece of information other than polarity. CT Information Question Constructions are formed through the use of CT Interrogative Contextual ObjRs. Some of these were shown already in Table 5.3 and they are repeated here in Table 7.1, along with the type of information that a speaker is requesting when they use each construction.

Interrogative	Requested Information
底啲 (ti <sup>11</sup> )tiaŋ <sup>55</sup> 'who'	person
乜个 mi? <sup>5</sup> kai <sup>55</sup> 'what'	thing
底块 (ti <sup>11</sup> )kɔ <sup>21</sup> 'where'	place
啲时 tiaŋ <sup>33</sup> si <sup>55</sup> 'when'	time
怎呢 tɕɔ <sup>52</sup> ni <sup>55</sup> how/why	manner, reason

Table 7.1: CT Interrogative Contextual ObjRs

Some CT Information Question Constructions are given in (30).

- (30) a. “怎呢            ∅    卡    危?”  
           “tɕɔ<sup>52</sup>ni<sup>55</sup>    ∅    k<sup>h</sup>aŋ<sup>2-5</sup>    kuŋ<sup>55</sup>?”  
           why            3SG    so    high  
           “‘Why [is it] so high?’” [G2F1US-20]
- b. “怎呢            跳跃        去?”  
           “tɕɔ<sup>52</sup>ni<sup>55</sup>    tiau<sup>21-53</sup>ue<sup>33</sup>    k<sup>h</sup>u<sup>21</sup>?”  
           how            jump        go  
           “‘How do [we] jump [over it]?’” [G2F1US-20]

- c. “阿妈 俺 怎呢 出 是来 着..  
 “a<sup>33</sup>-ma<sup>55</sup> naŋ<sup>52</sup> tɕɔ<sup>52</sup>ni<sup>55</sup> tɕ<sup>h</sup>uk<sup>2-5</sup> si<sup>24</sup> lai<sup>55</sup> tiɔ<sup>5</sup>..  
 NVOC-mom 1PL why exit be come RC..  
 抱 只 个 壳 在 俺 个 身 顶?”  
 p<sup>h</sup>ɔ<sup>24</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> k<sup>h</sup>ak<sup>2</sup> tɔ<sup>11</sup> naŋ<sup>52</sup> kai<sup>55-11</sup> siŋ<sup>33</sup> teŋ<sup>52</sup>?”  
 carry this LW shell LOC 1PL LW body atop  
 “Mom, why do we carry this shell on top of our body when we go out?””  
 [G2F1US-17]
- d. 只个 是 乜个  
 tɕi<sup>52</sup>kai<sup>55</sup> si<sup>11</sup> mi<sup>?</sup>kai<sup>55</sup>  
 PDCO COP what  
 ‘What is this?’ [G1M1-7]
- e. 只个 乜个  
 tɕi<sup>52</sup>kai<sup>55</sup> mi<sup>?</sup>kai<sup>55</sup>  
 PDCO what  
 ‘What is this?’ [G2M1-5, G1M1-7, G1F3-PS]

In CT Information Question Constructions, the CT Interrogative Contextual ObjR appears in the same position in the sentence where the referent would go, as evidenced by the pair of utterances in (31).

- (31) a. “汝 名 乜个?”  
 “lu<sup>52</sup> miã<sup>55-11</sup> mi<sup>?</sup>kai<sup>55</sup>?”  
 2SG name what  
 “What are you named?” [G2F1US-19]
- b. 伊 名 金鱼  
 i<sup>33</sup> miã<sup>55-11</sup> kim<sup>33</sup>hu<sup>55</sup>  
 3SG be.named gold-fish  
 ‘He is named Gold Fish.’ [G2F1US-20]

Because CT Information Question Constructions exhibit the same order as the canonical CT Declarative Constructions, and because the CT Interrogative Contextual ObjR

Constructions in Table 7.1 can also act as “Indefinite” Contextual ObjRs (see §5.2.2), there might be some confusion as to the interpretation of utterances with these words. CT Information Question Constructions can use the rising intonation strategy in order to distinguish between these two constructions, as in (32). With the rising intonation, the construction in (32) is a question (‘I see a woman doing what?’), not a declarative (‘I see a woman doing something.’).

- (32) Ø    睇    嫗    物    乜个  
 Ø    thoŋ<sup>52-24</sup> tɕa<sup>33</sup> bou<sup>52</sup>    mueŋ<sup>55-2</sup>    miŋ<sup>55</sup> kai<sup>55</sup>  
 1SG    see    woman    do    what  
 ‘[I] see [a] woman doing what?’ [G1M1-7]

This contrast is similarly showcased in (33), where the utterance in (33a) is spoken with the rising intonation strategy resulting in a CT Information Question Construction, while (33b) shows the same word 乜个 *miŋ<sup>55</sup> kai<sup>55</sup>* ‘what’ acting as a CT “Indefinite” Contextual ObjR Construction.

- (33) a.    有    乜个  
 u<sup>11</sup>    miŋ<sup>55</sup> kai<sup>55</sup>  
 have    what  
 ‘What is there? [G2M1-5, G2M1-6]
- b.    有    乜个  
 u<sup>11</sup>    miŋ<sup>55</sup> kai<sup>55</sup>  
 have    what  
 ‘There is something.’ [G2M1-5]

There was no evidence from Cambodian Teochew on the third type of questions, alternative questions, where the hearer is given two or more options to choose from (Croft,

forthcoming, p. 341).<sup>82</sup> Additionally, because the CT Interrogative Constructions shown in this section occurred in narrational texts, there was not an opportunity for a response from anyone. Therefore, CT Response Constructions are an area for future research.

### 7.1.3 CT Imperative-Hortative Constructions

It is generally agreed that an imperative construction is an appeal to the addressee to do something. Imperatives encompass all attempts to do this, whether it be through an order, request, command, invitation, warning, etc. (König and Siemund, 2007). There can also be appeals to the first or third person to do something (in English ‘Let’s go’ or ‘Let her go’, respectively). These are sometimes called “hortative” (Croft, forthcoming, p. 350, citing van der Auwera et al., 2003, pp. 49-52). Croft (forthcoming) follows van der Auwera et al. (2003) in adopting the term imperative-hortative constructions to encompass all appeals to do something, regardless of person. We will do the same here.

Given the nature of the data that was collected for this project, this type of speech act was not common in the primary texts. However, there were some examples in the secondary data, shown in (34).

- (34) a. “等 下 等 下 ... 孬 食 我. 放 我 落”  
 taŋ<sup>52</sup> e<sup>24</sup> taŋ<sup>52</sup> e<sup>24</sup> ... mo<sup>11</sup> tɕiaɪ<sup>ʔ5-2</sup> ua<sup>52</sup>. paŋ<sup>21-52</sup> ua<sup>52</sup> lo<sup>ʔ5</sup>”  
 wait below wait below ... NEG.can eat 1SG put 1SG down  
 “‘Wait a second, wait a second! ... Don’t eat me! Put me down!’”  
 [G2F1US-18]

---

<sup>82</sup> But see Xu (2007, pp. 265-8) and Low (2014, p. 47) for Jieyang and Singapore Alternative Question Constructions.

- b. “汝 等 睇 我 跳 起 去”  
 “lu<sup>52</sup> taŋ<sup>52</sup> tʰɔ̃<sup>52-24</sup> ua<sup>52</sup> tiau<sup>21-53</sup> kʰi<sup>52</sup> kʰu<sup>21</sup>”  
 2SG wait see 1SG jump up go  
 ‘Wait and watch me jump over.’ [G2F1US-20]
- c. “俺 去 无?”  
 “naŋ<sup>52</sup> kʰu<sup>21</sup> bɔ̃<sup>55?</sup>”  
 1PL go NEG  
 ‘‘Let’s go, no?’’ [G2F1US-20]
- d. 伊 但 “俺 泅 起 去”  
 i<sup>33</sup> tã<sup>21</sup> “naŋ<sup>52</sup> siu<sup>55-11</sup> kʰi<sup>52</sup> kʰu<sup>21</sup>”  
 3SG say 1PL swim up go  
 ‘He says, ‘‘Let’s swim up.’’ [G2F1US-20]

It was shown in Chapter 5 that the use of overt reference in Cambodian Teochew is optional if the referent can be recovered from the context. This is also the case with CT Imperative-Hortative Constructions, where the use of 汝 *lu*<sup>52</sup> ‘2SG’ is optional, as shown in the complementary examples in (35).

- (35) a. 伊 但 “□! 汝 睇”  
 i<sup>33</sup> tã<sup>21-53</sup> “ɔ̃<sup>33!</sup> lu<sup>52</sup> tɔ̃<sup>52</sup>”  
 3SG say EXLA 2SG look  
 ‘He says, ‘‘Oh! Look!’’ [G2F1US-18]
- b. 伊 但 “睇”  
 i<sup>33</sup> tã<sup>21-53</sup> “tʰɔ̃<sup>52</sup>”  
 3SG say look  
 He says, ‘‘Look!’’ [G2F1US-18]

Since all attempts to get someone to do something are considered CT Imperative-Hortative Constructions, included in this type of construction are examples like (36) which express deontic modality with the predicate 着 *tiɔ̃*<sup>ɔ̃</sup> ‘must’.

- (36) a. “俺 卡 细. 好.. 着 觅 个 空 啊”  
 “nan<sup>52</sup> k<sup>h</sup>aɿ<sup>2-5</sup> soi<sup>21</sup>. hɔ<sup>52</sup>.. tiɔɿ<sup>5-2</sup> tɕ<sup>h</sup>ue<sup>11</sup> kai<sup>55-11</sup> k<sup>h</sup>aŋ<sup>33</sup> a<sup>33</sup>”  
 1PL too small. good.. must look.for IK hole PRT  
 “We are too small. [We] must look for the hole.” [G2F1US-20]
- b. “俺.. 俺 着 汹 起 去”  
 “nan<sup>52</sup>..nan<sup>52</sup> tiɔɿ<sup>5-2</sup> siu<sup>55-11</sup> k<sup>h</sup>i<sup>52</sup> k<sup>h</sup>u<sup>21</sup>”  
 1PL.. 1PL must swim up go  
 “We must swim up.” [G2F1US-20]

The negative of the imperative-hortative construction is a prohibitive construction. The CT Prohibitive Construction uses the prohibitive negator strategy in that a special negative form, 孬 mɔ<sup>11</sup> ‘NEG.can’ is used only in this construction, as in (37).

- (37) a. “孬 食 我”  
 mɔ<sup>11</sup> tɕiaɿ<sup>5-2</sup> ua<sup>52</sup>  
 NEG.can eat 1SG  
 “Don’t eat me!” [G2F1US-18]
- b. 妈妈 伊 但 “OK 孬 哭 了”  
 ma<sup>55-11</sup>ma<sup>55</sup> i<sup>33</sup> tã<sup>21-53</sup> “OK mɔ<sup>52-24</sup> k<sup>h</sup>au<sup>21-53</sup> ɔu<sup>11</sup>”  
 mama 3SG say “OK NEG.can cry already”  
 ‘Mama, she says, “Ok, stop crying.” [G2F1US-17]
- c. “孬 哭 个 团”  
 “mɔ<sup>52-24</sup> k<sup>h</sup>au<sup>21</sup> kai<sup>55-11</sup> kiã<sup>52</sup>”  
 NEG.can cry IK child  
 “Don’t cry child.” [G2F1US-17]

The pair of examples in (38) shows the contrast between the same predicate being used in a CT Prohibitive Construction (38a) and in a CT Declarative Negation Construction (38b).

- (38) a. “个 花 汝 姦 摘 伊 口”  
 “kai<sup>55-11</sup> hue<sup>33</sup> lu<sup>52</sup> mo<sup>52-24</sup> tia<sup>ʔ2-5</sup> i<sup>33</sup> nia<sup>11</sup>”  
 “CL flower 2SG NEG.can pick 3SG PRT”  
 “The flower, don’t pick it.’ [G2F1US-19]
- b. 汝 无 摘 伊 真好  
 lu<sup>52</sup> bo<sup>55-11</sup> tia<sup>ʔ2-5</sup> i<sup>33</sup> teiŋ<sup>33</sup> ho<sup>52</sup>  
 2SG NEG pick 3SG very good  
 “It’s very good that you didn’t pluck it.” [G2F1US-19]

#### 7.1.4 CT Complex Predicate Constructions

According to Croft (forthcoming, p. 369), complex predicate constructions are made up of two or more concepts that are combined into one predication. The information packaging is predication and the morphosyntactic form is complex. What exactly constitutes a complex predicate, however, can be hard to define, as complex predicates can become simple predicates due to grammaticalization and lexicalization (Croft, forthcoming, p. 369).

For example, recall that the predicate 咀 *tã<sup>21</sup>* ‘say’ has grammaticalized into use a complementizer, yet still has retained its lexical usages. When combined with speech act predicates like 问 *muŋ<sup>11</sup>* ‘ask’, there can be ambiguity between whether the clause in an example like (39) constitutes a CT Eventive Complex Predicate Construction, with a meaning like ‘He asks saying...’ or if the 咀 *tã<sup>21</sup>* is acting like a grammaticalized complementizer and thus the construction just contains the simple predicate 问 *muŋ<sup>11</sup>* ‘ask’.

- (39) 伊 问 咀 ...  
 i<sup>33</sup> muŋ<sup>11</sup> tã<sup>21</sup> ...  
 3SG ask say/COMP ...  
 ‘He asks saying ...’  
 ‘He asks that...’

Of the two or more concepts that make up a complex predicate, one of them is usually an action, or a “process in which change occurs” (Croft, forthcoming, p. 370). When the other component is another action/process, the construction is called an eventive complex predicate (§7.1.4.1). When the other component denotes a state, Croft (forthcoming) calls these stative complex predicate constructions (§7.1.4.2).

#### *7.1.4.1 CT Eventive Complex Predicate Constructions*

We have shown complex sentence constructions such as CT Action Reference Constructions and CT Action Modification Constructions in which two or more clauses combine in a single utterance (see also §7.1.4 below). Eventive complex predicate constructions differ from such complex sentences in that the components of the predicate are describing subevents of one singular event (Croft, forthcoming, p. 371). Croft (forthcoming, pp. 372-3) defines a “single” event as one where the combination of events: shares the same tense, aspect, and modality values; constitutes a single argument structure with one configuration of participants; and constitutes a single assertion, where one part cannot be negated without negating the other.

The contrast between an event with a single argument structure and one without is illustrated in the pair of sentences in (40). Both sentences have a null argument that is the agent performing the action of 拖 *t<sup>h</sup>ua*<sup>33</sup> ‘pull’. In (40a), the patient of this first predicate is 狗 *kau*<sup>52</sup> ‘dog’, which is also the agent of the second predicate, 行 *kiā*<sup>55</sup> ‘walk’. That the referent 狗 *kau*<sup>52</sup> ‘dog’ has a different semantic role for each predicate excludes this type of construction from being a CT Eventive Complex Predicate Construction. The same cannot be said of (40b),

because 布 *pɔu*<sup>21</sup> ‘linen’ does not have two semantic roles. Therefore (40b) is a CT Eventive Complex Predicate Construction that uses the serial predicate strategy.<sup>83</sup>

- (40) a. 只个 Ø 拖 狗 行  
*tɕei*<sup>55</sup> Ø *t<sup>h</sup>ua*<sup>33</sup> *kau*<sup>52</sup> *kiã*<sup>55</sup>  
 PDCO 3SG pull dog walk  
 ‘[In] this (one [the picture]), [someone] takes a dog for a walk.’  
 [G1F1-10]
- b. Ø 拖 布 行  
 Ø *t<sup>h</sup>ua*<sup>33</sup> *pɔu*<sup>21</sup> *kiã*<sup>55</sup>  
 3PL pull linen walk  
 ‘[They] pull the linen walking.’ [G1F1-10]

(41) is an example of a CT Eventive Complex Predicate Construction, where the 去放屎 *k<sup>h</sup>ur*<sup>21-53</sup> *paŋ*<sup>52</sup> *sai*<sup>52</sup> ‘go poop’ makes up the complex predicate as a motion event (see below), not the 行 *kiã*<sup>55</sup> ‘walk’<sup>84</sup> and 去 *k<sup>h</sup>ur*<sup>21</sup> ‘go’. This example also uses the serial predicate strategy.

- (41) Ø 行 狗团 去 放屎 了  
 Ø *kiã*<sup>55-11</sup> *kau*<sup>52-24</sup> *kiã*<sup>52</sup> *k<sup>h</sup>ur*<sup>21-53</sup> *paŋ*<sup>52</sup> *sai*<sup>52</sup> *ɔu*<sup>11</sup>  
 3SG walk puppy go poop already  
 ‘[She] walked the dog to go poop.’ [G1M1-7]

<sup>83</sup> This is the “serial verb strategy” in Croft (forthcoming, p. 376). It is frequently called the “serial verb construction” or SVC in the literature (Aikhenvald & Dixon, 2006; Haspelmath, 2016; Li & Thompson 1981, pp. 594-620; Matthews, 2006; Self, 2014). See Paul (2008) for a critique of the use of the term “construction” in this context.

<sup>84</sup> R.J. LaPolla (personal communication, December 27, 2021) noted that the use of 行 *kiã*<sup>55</sup> ‘walk’ in this construction is interesting from a Chinese point of view, because we would usually expect a different predicate, as in Mandarin’s 遛狗 *liú gǒu* ‘to walk a dog’. Indeed, this pairing is found in the Teochew dictionary TCKnow LLC. (2015), 遛狗 *liu*<sup>52</sup> *kau*<sup>52</sup> ‘to walk a dog’. There were no tokens of the use of the predicate 遛 *liu*<sup>21</sup> ‘walk’ in the current data set on Cambodian Teochew. Future research should look at the usages of these two predicates in this context.

The serial predicate strategy is actually made up of a collection of strategies as classified by Durie (1997) related to whether the predicates are contiguous, or not contiguous, and a single word or separate words, and where the locus of inflection is. Since there is no inflection on CT Predicates, we are primarily concerned with the first two characterizations.

In the prototypical serial predicate strategy, the predicates are contiguous and separate words (Croft, forthcoming, p. 377). This is illustrated in the examples of CT Eventive Complex Predicate Constructions using the serial predicate strategy in (42). It is very productive in Cambodian Teochew to have a posture predicate such as 坐 *tɕɔ<sup>24</sup>* ‘sit’ or 企 *k<sup>h</sup>ia<sup>24</sup>* ‘stand’ combine with an activity predicate to create a CT Eventive Complex Predicate Construction, as in (42).<sup>85</sup> Given that the predicates in this construction are construed as one event, it is generally the case that all but the last predicate undergo tone sandhi, though there are exceptions. Future research could explore this issue further.

- (42) a. 只 个 父 只 个 团 坐 食 碱酸  
*tɕei<sup>52</sup> kai<sup>55-11</sup> pe<sup>24</sup> tɕei<sup>52</sup> kai<sup>55-11</sup> kiã<sup>52</sup> tɕɔ<sup>24-11</sup> tɕial<sup>5-2</sup> kiam<sup>11</sup> sunj<sup>33</sup>*  
 this LW father this LW child sit eat food  
 ‘This father and this child sit, eating food.’ [G3F1-3]
- b. 只个 伊 坐 睇 书  
*tɕei<sup>55</sup> i<sup>33</sup> tɕɔ<sup>24-11</sup> t<sup>h</sup>ʃi<sup>52-24</sup> tɕu<sup>33</sup>*  
 PDCO 3SG sit look book  
 ‘[In] this (one [the picture]), he sits reading a book.’ [G1F1-12]

<sup>85</sup> See Matthews (2006, pp. 82-3) for a similar pairing in Cantonese.

- c. 只 隻 鸡翁 企 在 篱笆 顶  
 tɕei<sup>52</sup> tɕiaŋ<sup>2-5</sup> kɔi<sup>33</sup>an<sup>11</sup> k<sup>h</sup>ia<sup>24</sup> tɔ<sup>11</sup> li<sup>11</sup>pa<sup>33</sup> teŋ<sup>52</sup>  
 this CLF.ANIMAL rooster stand LOC fence atop  
 孛团 在 企 睇 Ø  
 nɔu<sup>33</sup>kiã<sup>52</sup> tɔ<sup>11</sup> k<sup>h</sup>ia<sup>24-11</sup> t<sup>h</sup>ɔi<sup>52</sup> Ø  
 kid PROG stand see 3SG  
 ‘This rooster is standing on top of a fence and [some] kids are standing looking [at it].’ [G1F3-15]

The two (or more) predicates do not have to be contiguous in CT Eventive Complex Predicate Constructions. They can still be construed as one event even if there is an intervening argument, as in (43). This was also shown in the example in (40b).

- (43) a. 只个 坐 飞机 睇 环境  
 tɕei<sup>53</sup> tɕɔ<sup>24-11</sup> pue<sup>33</sup>ki<sup>33</sup> t<sup>h</sup>ɔi<sup>52-24</sup> huan<sup>11</sup>keŋ<sup>52</sup>  
 PDCO sit plane see surroundings  
 ‘[In] this (one [the picture]), he sits on a plane looking at the surroundings.’ [G1F1-8]
- b. 只个 孃妣... 口 位沙 坐 睇 电视 响\*薰  
 tɕei<sup>53</sup> tɕa<sup>33</sup>bou<sup>33</sup>...na<sup>11</sup> ui<sup>11</sup>-sua<sup>33</sup> tɕɔ<sup>24-11</sup> t<sup>h</sup>ɔi<sup>52-11</sup> tiaŋ<sup>11</sup>si<sup>24</sup> kuŋ<sup>5</sup>hun<sup>33</sup>  
 this.LW woman... be.at place-sand sit watch TV smoke  
 ‘This woman... is in the desert, sitting, watching TV, and smoking.’  
 [G3F1-1]

CT Eventive Complex Predicate Constructions are also highly productive for motion events, as in (44).

- (44) Ø 行 入.. 客厅 ...  
 Ø kia<sup>55-11</sup> dzip<sup>5-2</sup>..k<sup>h</sup>eɿ<sup>2-5</sup>-t<sup>h</sup>uŋ<sup>55</sup> ...  
 3SG walk enter.. guest-room ...  
 ‘[She] walks into the living room...’ [G3F1-2]

We’ll use Croft’s (forthcoming, p. 418) summary of Talmy’s (1972) decomposition of a motion event to analyze the construction in (44), which is made up of five parts, in (45).

- (45) (i) the *motion* itself  
 (ii) the *manner* of motion (行 kia<sup>55</sup> ‘walk’)  
 (iii) the *path* of motion (from outside of the living room into it, 入 dzip<sup>5</sup> ‘enter’)  
 (iv) the *figure*, or the object that is moving (the null 3SG)  
 (v) the *ground*, or an object acting as a reference point for the path of motion (客厅 k<sup>h</sup>eɿ<sup>2-5</sup>-t<sup>h</sup>uŋ<sup>55</sup> living room)

Thus we can see that in (44), the first predicate expresses the manner of motion, while the second predicate expresses the path of motion, and the figure and ground are represented by arguments of the predicate.

Similarly, in another CT Eventive Complex Predicate Construction using the serial predicate strategy in (46), the manner of motion is 行 kia<sup>55</sup> ‘walk’, the figure is 孛囡 nɔu<sup>33</sup>kiã<sup>52</sup> ‘child’, the ground is 读书 t<sup>h</sup>ak<sup>5-2</sup>tɕu<sup>33</sup> ‘school’, and the path is the deictic direction 去 k<sup>h</sup>u<sup>21</sup> ‘go’.

- (46) 孛囡 行 去 读书  
 nɔu<sup>33</sup>kiã<sup>52</sup> kia<sup>55-11</sup> k<sup>h</sup>u<sup>21-53</sup>t<sup>h</sup>ak<sup>5-2</sup>tɕu<sup>33</sup>  
 child walk go read-book  
 ‘A child is walking to school.’ [G1F1-12]

Frequently used manner incorporating predicates in the current data set on Cambodian Teochew include 行 kiã<sup>55</sup> ‘walk’, 飞 pue<sup>33</sup> ‘fly’, 泅 siu<sup>55</sup> ‘swim’, and 爬 pe<sup>55</sup> ‘crawl’. Path of

motion predicates include 出 *tɕ<sup>h</sup>uk<sup>2</sup>* ‘exit’, 入 *dzip<sup>5</sup>* ‘enter’, 跋 *pua<sup>5</sup>* ‘fall’, 落 *lo<sup>5</sup>* ‘descend’, 起 *k<sup>h</sup>i<sup>52</sup>* ‘ascend’, and 过 *kue<sup>21</sup>* ‘cross’.<sup>86</sup> Deictic direction predicates include 来 *lai<sup>55</sup>* ‘come’ and 去 *k<sup>h</sup>u<sup>21</sup>* ‘go’.

The order in the constructions in (44) and (46) was figure + manner + path + ground, but this order is not obligatory. Motion events do not even have to express all of the five semantic components in (45) (Croft, forthcoming, p. 420). The example in (47) shows a CT Eventive Complex Predicate Construction with the order figure + path + manner. No ground is expressed in this example. The example in (48) only has a figure + deictic motion + activity. There is no manner or ground.

- (47) 有 飞机 落 飞  
 u<sup>11</sup> pue<sup>33</sup>ki<sup>33</sup> lo<sup>5-2</sup> pue<sup>33</sup>  
 have plane descend fly  
 ‘There is a plane landing.’ [G1F4-16]

- (48) Ø 咁 伊 个 翁人 恬 去 食 酒  
 Ø tā<sup>21-53</sup> i<sup>33</sup> kai<sup>55-11</sup> aŋ<sup>33</sup>naŋ<sup>55</sup> tiam<sup>11</sup> k<sup>h</sup>u<sup>21-53</sup> tɕia<sup>5-2</sup> tɕiu<sup>52</sup>  
 3SG say 3SG LW husband always go drink alcohol  
 合 伊 个 朋友  
 ka<sup>2-5</sup> i<sup>33</sup> kai<sup>55-11</sup> p<sup>h</sup>ej<sup>55-11</sup>iu<sup>52</sup>  
 with 3SG LW friend  
 ‘[She] says her husband always goes to drink alcohol with his friend(s).’  
 [G3F2-4]

There can be several path of motion predicates in one construction, as in (49) which shows the order figure + path + path + deictic motion.

<sup>86</sup> See Matthews (2006, p. 83) for the use of a similar set of predicates in motion events in Cantonese.

- (49) 猫 跋 落 来 啊 口  
 ŋiau<sup>33</sup> pua<sup>ʔ5</sup> lo<sup>ʔ5-2</sup> lai<sup>55-11</sup> a<sup>11</sup> no<sup>11</sup>  
 cat fall descend come PRT PRT  
 ‘[The] cat falls down.’ [G1F2-14]

The difference between the deictic direction predicates can be explained with the examples in (50). While they both have the order figure + path + path + deictic + path + ground + path, (50a) uses the deictic 来 *lai*<sup>55</sup> ‘come’ while (50b) uses the deictic 去 *k<sup>h</sup>u*<sup>21</sup> ‘go’. We show the context for (50a) in Figure 7.1, and the context for (50b) in Figure 7.2.

- (50) a. ... 就 伊.. 跋 落 来 入.. 泥 底畔  
 ... tɕiu<sup>11</sup> i<sup>33</sup>.. pua<sup>ʔ5</sup> lo<sup>ʔ5</sup> lai<sup>55</sup> dzip<sup>5-2</sup>.. p<sup>h</sup>ɔk<sup>5</sup> tɔi<sup>52-11</sup> pa<sup>ʔ55</sup>  
 ... then 3SG.. fall descend come enter.. mud inside  
 ‘...then he falls down into the mud.’ [G3F1-3]
- b. 只个 孛团 跋 落 去 入 水水 底  
 tɕei<sup>55</sup> nou<sup>33</sup>kiã<sup>52</sup> pua<sup>ʔ5</sup> lo<sup>ʔ5</sup> k<sup>h</sup>u<sup>21</sup> dzip<sup>2</sup> tɕu- tɕui<sup>52-24</sup> tɔi<sup>52</sup>  
 PDCO child fall descend go enter water water inside  
 ‘[In] this (one [the picture]), the child falls into the water.’ [G1F1-11]

In Figure 7.1, the point that represents the ground from which the speaker is talking is the bottom of the wall. When the cat falls into the mud at the bottom of the wall, the cat is falling towards the speaker, so 来 *lai*<sup>55</sup> ‘come’ is used. Conversely, in Figure 7.2, the point that represents the ground for the speaker is where the deer is standing. Thus when the child falls into the water, he is falling away from the speaker, so 去 *k<sup>h</sup>u*<sup>21</sup> ‘go’ is used. Because of this distinction, it is not expected that there would be multiple deictic direction predicates in one CT Eventive Complex Predicate Construction. Similarly, it would lend a strange semantic interpretation to have more than one manner incorporating predicate in one construction. Indeed, neither of these combinations were found in the data.



Figure 7.1: The Cat Story context for example (50a)

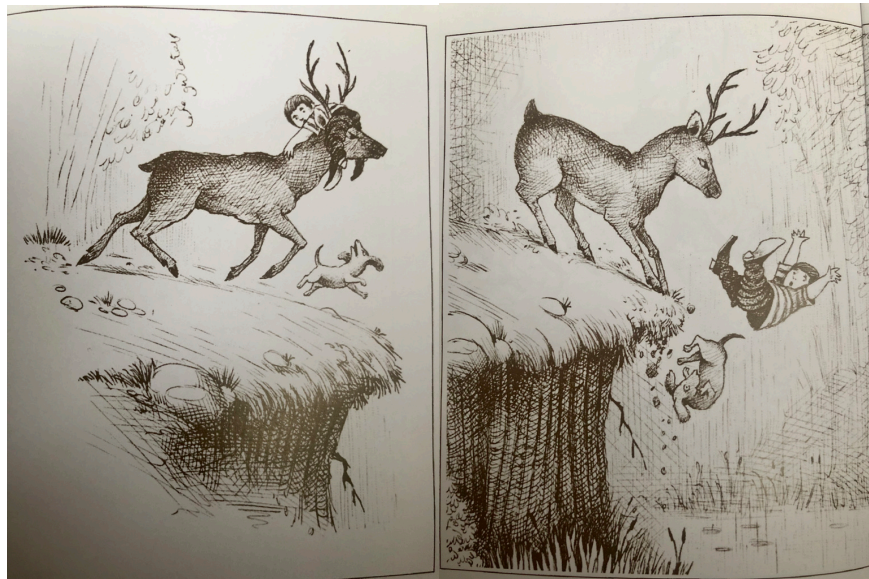


Figure 7.2: The Frog Story context for example (50b)

The examples in (50) both used an additional non-predicate path word 底畔/底  $tɔ̃^{52}$ - $^{11}pã^{55}/tɔ̃^{52}$  ‘inside’ to emphasize the path of the motion. This word does not seem to be required, as the example in (51) with a similar meaning does not contain such a word, nor do the examples in (44) and (46-49).

- (51) 伊 跋 入 水 好 ...  
 $i^{33}$   $puaŋ^2$   $jip^2$   $tɕiu^{52}$   $hɔ^{52}$  ...  
 3SG fall into water good ...  
 ‘He has fallen into the water...’ [G3F1-3]

Optionality is further shown in examples like (52) where the same complex predicate 飞起  $pue^{33} k^h i^{52}$  ‘fly up’ can be used with or without 顶畔  $teŋ^{52-24} pã^{55}$  ‘atop’, in (52a). Similarly, we can compare the last utterance in (52a) with the utterance in (52b) where the same complex predicate 飞起去  $pue^{33} k^h i^{52} k^h u^{21}$  ‘fly up go’ can again be used with or without 顶畔  $teŋ^{52-24} pã^{55}$  ‘atop’.

- (52) a. 伊 飞 起 顶畔.            Ø    飞 起.  
 $i^{33}$   $pue^{33}$   $k^h i^{52-24}$   $teŋ^{52-24} pã^{55}$ .    Ø     $pue^{33}$   $k^h i^{52}$ .  
 3SG fly up atop            3SG fly up  
 伊 飞 起 去 了  
 $i^{33}$   $pue^{33}$   $k^h i^{52}$   $k^h u^{21}$   $ɔu^{11}$   
 3SG fly up go already  
 ‘She flies up above. [She] flies up. She has flown up.’ [G1F1-12]
- b. 伊 有 翼 就 会- 会会 变成  
 $i^{33}$   $u^{11}$   $sik^5$   $tɕiu^{11}$   $hui-$   $ɔi^{11}$   $ɔi^{11}$   $pian^{21-53} tɕiã^{55}$   
 3SG have wing then can- can RED change  
 飞 起 去 顶畔  
 $pue^{33}$   $k^h i^{52}$   $k^h u^{21}$   $teŋ^{52-24} pã^{55}$   
 fly up go atop  
 ‘It [the butterfly] has wings so it can change and fly up above.’  
 [G2F1US-17]

While the motion event in (53) is not a complex predicate, it further shows the optionality of a path word like 底畔 *tɔi<sup>52-24</sup>paɪ<sup>55</sup>* ‘inside’ with the path of motion predicate 落 *lɔʔ<sup>6</sup>* ‘fall’.

- (53) 就 落 水. 落 水 底畔 就 睇 着 鱼  
*tɕiu<sup>11</sup> lɔʔ<sup>5</sup> tɕui<sup>52</sup>. lɔʔ<sup>5-2</sup> tɕui<sup>52</sup> tɔi<sup>52-24</sup>paɪ<sup>55</sup> tɕiu<sup>11</sup> θɔɪ<sup>52-24</sup> tɔʔ<sup>2</sup> hu<sup>55</sup>*  
 then fall water fall water inside then see RC fish  
 ‘Then [it] falls in the water. After falling in the water, [it] saw fish.’ [G1F1-9]

For one path of motion predicate 出 *tɕ<sup>h</sup>uk<sup>2</sup>* ‘exit’, evidence points to the use of word order to show the path of motion. Thus, the ground when it occurs before the predicate acts as the point of reference from which the motion begins while the ground that appears after the predicate marks the point where the motion ends. In this way, the word order is iconic. The examples in (54) show when the ground is the starting point (appearing before the predicate) while the examples in (55) show when the ground is the end point (appearing after the predicate).

- (54) a. 人 在 口, 人 读书 出 去 了  
*naŋ<sup>55-11</sup> tɔ<sup>11</sup> khau<sup>52-24</sup>, naŋ<sup>55-11</sup> t<sup>h</sup>ak<sup>5-2</sup>tɕu<sup>33</sup> tɕ<sup>h</sup>uk<sup>2</sup> k<sup>h</sup>u<sup>21</sup> ɔu<sup>11</sup>*  
 person be.at outside person school exit to go already  
 ‘People are outside, people have gone out of the school.’ [G1F1-12]
- b. 𠵼 只个 孳团 伊 读书 出 来 ...  
*hai<sup>33</sup> tɕia<sup>53</sup> nɔu<sup>33</sup>kiã<sup>52</sup> i<sup>33</sup> t<sup>h</sup>ak<sup>5-2</sup>tɕu<sup>33</sup> tɕ<sup>h</sup>uk<sup>2</sup> lai<sup>55</sup> ...*  
 and this.LW child 3SG school exit come ...  
 ‘And this child, she leaves school...’ [G1F1-12]

- (55) a. Ø 行 出 口 迺迺  
 Ø kiã<sup>55-11</sup> tɕ<sup>h</sup>uk<sup>2-5</sup> k<sup>h</sup>au<sup>52</sup> t<sup>h</sup>ik<sup>5</sup>t<sup>h</sup>ɔ<sup>55</sup>  
 3SG walk out outside stroll  
 ‘[She] walks outside for a stroll.’ [G1F4-16]
- b. Ø 行 出 外口  
 Ø kiã<sup>55-11</sup> tɕ<sup>h</sup>uk<sup>2</sup> wua<sup>11</sup>k<sup>h</sup>au<sup>52</sup>  
 3SG walk exit outside  
 ‘[She] walks outside.’ [G1F1-12]
- c. 只个 Ø 来 出 外口  
 tɕei<sup>55</sup> Ø lai<sup>55-11</sup> tɕ<sup>h</sup>uk<sup>2</sup> bua<sup>11</sup>k<sup>h</sup>au<sup>52</sup>  
 PDCO 3PL come exit outside<sup>87</sup>  
 ‘[In] this (one [the picture], they come outside.’ [G1F1-11]
- d. 伊 爬 出 来 口  
 i<sup>33</sup> pe<sup>33</sup> tɕ<sup>h</sup>uk<sup>2</sup> lai<sup>55-11</sup> k<sup>h</sup>au<sup>52</sup>  
 3SG crawl exit come outside  
 ‘He crawls outside.’ [G1F2-13]

There were instances where speaker G3F1 did not follow this iconicity pattern, as in (56). The reference point in each of these utterances is the starting point of the motion, though it appears after the predicate 出 tɕ<sup>h</sup>uk<sup>2</sup> ‘exit’. However, in these examples, the Khmer loanword 𑜀𑜢𑜤𑜰𑜫 pi<sup>33</sup> ‘from’ is used to indicate that the ground is the starting point. Thus it might be that the starting point is allowed to occur after the predicate 出 tɕ<sup>h</sup>uk<sup>2</sup> ‘exit’ if an additional morpheme such as 𑜀𑜢𑜤𑜰𑜫 pi<sup>33</sup> ‘from’ is used to disambiguate. More CT Eventive Complex Predicate Constructions are needed in order to confirm the patterns noted here.

---

<sup>87</sup> This word appears to show a /g/ → /w/ → /b/ sound change for speaker G1F1. This could be the pathway for the /g/ → /b/ sound change that was noted in ‘moon’ for the same speaker.

- (56) a. Ø 飞 出.. 从 只个 海  
 Ø pue<sup>33</sup> tɕ<sup>h</sup>uk<sup>2-5</sup>..pi<sup>33</sup> tɕia<sup>53</sup> hai<sup>52</sup>  
 3SG fly exit.. from this.LWsea  
 ‘[The plane] flies away from this sea.’ [G3F1-1]
- b. 就 伊 爬\* 出 从 伊 个 褥  
 tɕiu<sup>31</sup> i<sup>33</sup> peŋ<sup>5</sup> tɕ<sup>h</sup>uk<sup>2-5</sup> pi<sup>33</sup> i<sup>33</sup> kai<sup>55-11</sup> dʒək<sup>5</sup>  
 then 3SG climb exit from 3SG LW mattress  
 ‘Then he climbs out from his crib.’ [G3F1-2]
- c. 就 伊 爬 出 从 房  
 tɕiu<sup>11</sup> i<sup>33</sup> pe<sup>55-11</sup> tɕ<sup>h</sup>uk<sup>2-5</sup> pi<sup>33</sup> paŋ<sup>55</sup>  
 then 3SG crawl exit from room  
 ‘Then he crawls out from his room.’ [G3F1-2]

#### 7.1.4.2 CT Stative Complex Predicate Constructions

In stative complex predicate constructions, one of the concepts in the complex predicate is an action, while the other concept denotes a state (Croft, forthcoming, p. 371). There were not many examples of CT Stative Complex Predicate Constructions in the data, though they are explored here.

In depictive complex predicate constructions, the property predication part of the complex predicate is participant oriented because it is describing a characteristic of one of the participants (Croft, forthcoming, p. 401). CT Depictive Complex Predicate Constructions using the serial predicate strategy are given in (57).

- (57) a. 伊人 跳舞 清心 绝  
 i<sup>33</sup>naŋ<sup>11</sup> t<sup>h</sup>iau<sup>21-53</sup>bu<sup>52</sup> tɕ<sup>h</sup>eŋ<sup>33</sup>sim<sup>33</sup> tɕɔŋ<sup>5</sup>  
 3PL dance content very  
 ‘They dance very happily.’ [G1F4-16]

- b. 伊 食 好 就 伊 欢喜 死  
 i<sup>33</sup> tɕiaŋ<sup>5-2</sup> hɔ<sup>52</sup> tɕiu<sup>11</sup> i<sup>33</sup> huã<sup>33</sup>hi<sup>52</sup> ʔ<sup>24</sup>  
 3SG eat good then 3SG happy sleep  
 ‘He has eaten then he happily sleeps.’ [G3F1-3]

In manner complex predicate constructions, the property predication part of the complex predicate is event oriented because it is describing a property of the event (Croft, forthcoming, p. 401). CT Manner Complex Predicate Constructions using the serial predicate strategy are given in (58).

- (58) a. 飞 危  
 pue<sup>33</sup> kuĩ<sup>11</sup>  
 fly high  
 ‘[It] flies high.’ [G3F1-1]
- b. Ø 爬 歪 猛  
 Ø pe<sup>55</sup> bɔi<sup>11</sup> mɛ<sup>52</sup>  
 1PL crawl NEG.can fast  
 ‘[We] don’t crawl fast. [Literally: We crawl cannot fast.]’ [G2F1US-17]
- c. Ø 爬 卡\* 慢慢  
 Ø pe<sup>55</sup> k<sup>h</sup>aŋ<sup>2-5</sup> maŋ<sup>11</sup>maŋ<sup>11</sup>  
 1PL crawl too slow RED  
 ‘[We] crawl too slowly.’ [G2F1US-17]
- d. 伊人 跳舞 好, 好 绝  
 i<sup>33</sup>naŋ<sup>11</sup> t<sup>h</sup>iau<sup>21-53</sup>bu<sup>52</sup> hɔ<sup>52</sup>, hɔ<sup>52-24</sup> tɕɔŋ<sup>5</sup>  
 3PL dance good, good very  
 ‘They dance well, very well.’ [G1F4-16]

Given the relatively few examples of this construction in the current data set, we cannot say much about their behavior. Future research could investigate things like the ordering of the property predication and action predication, the use of admodifiers such as 绝 tɕɔŋ<sup>5</sup> ‘very’ and 卡\* k<sup>h</sup>aŋ<sup>2</sup> ‘too’, and the use of negation.

### 7.1.5 Complex Sentences

Per Croft (forthcoming, p. 429), complex sentences are constructions that consist of two or more clausal or clause-like constructions. We have already covered two types of complex sentences in previous Chapters: action reference, where a clause is acting as an argument of another clause, and action modification, where a clause is acting as a modifier of a referent in another clause.

Complex sentences with temporal and other relations between events can be expressed via coordinate clause constructions (§7.1.5.1) and/or adverbial clause constructions. Temporal and causal relations between events describe a semantic relationship, whereas coordination and subordination encode information packaging. Coordinate clause constructions represent a complex figure where two events are part of a complex whole (Croft, forthcoming, p. 435, citing Wierzbicka, 1980). Conversely, adverbial clause constructions represent a figure ground information packaging, where there is an asymmetrical relation between the events (Croft, forthcoming, p. 434).

Other types of complex sentences include comparative, equative (§7.1.5.2), conditional (§7.1.5.3), and concessive constructions.

Given the nature of the collected data in this study, the utterances collected tended to have more simple rather than complex constructions. Thus, there are many types of complex constructions highlighted by Croft (forthcoming) that were either not found in the current data set, or not sufficiently represented, including adverbial clause constructions, comparative constructions,<sup>88</sup> and concessive constructions. Future research should attempt to procure

---

<sup>88</sup> See Xu (2007)'s Chapter 11 on Jieyang Teochew Comparative Constructions and Low (2014, p. 30) for Singapore Teochew Comparative Constructions.

different types of naturalistic data to see to the extent and usage of constructions of this nature in Cambodian Teochew.

#### 7.1.5.1 CT Coordinate Clause Constructions (and CT Coordinate Constructions)

Coordinate clause constructions are a subtype of coordinate constructions (Croft, forthcoming, p. 436). We will look at the coordination of both referents and of clauses in this section, using the term COORDINAND (Haspelmath, 2004) to refer to one of the items being conjoined.

Coordinate constructions are commonly subcategorized into three types: conjunctive coordination constructions, disjunctive coordination constructions, and adversative coordination constructions (Croft, forthcoming, p. 436).

Conjunctive coordination constructions perform the additive function. CT Conjunctive Coordination Constructions with two or more referents can use the syndetic strategy, where an overt morpheme, the conjunction 合 *kaɿ<sup>2</sup>* ‘and’, codes the coordination relationship between the two or more coordinands. Haspelmath uses the term MONOSYNDETIC coordination to describe the strategy when there are fewer coordinators than coordinands.<sup>89</sup> This is shown in (59).

---

<sup>89</sup> Low (2014, p. 29) showed the use of a bisyndetic strategy for Singapore Teochew Conjunctive Coordinate Clause Constructions, where there is an equal number of coordinators and coordinands, using 又 *iu<sup>1</sup>*, as in (i). The use of this conjunction was not found in the current data set on Cambodian Teochew.

- (i)    liou<sup>53</sup>   ua<sup>53</sup>   iu<sup>11</sup>   k<sup>h</sup>i<sup>11</sup>   iu<sup>33</sup>   tsij<sup>33</sup>   kan<sup>35</sup>   tʃiɔŋ<sup>55</sup>  
       PFV    1SG    SIM    angry   SIM    very    anxious  
       ‘Then I was both angry and anxious.’ (Low, 2014, p. 29)

- (59) a. 鹿 合 老猴 啊是 七个  
tek<sup>5</sup> kaɿ<sup>2</sup> lau<sup>24-11</sup>kau<sup>55</sup> a<sup>11</sup>si<sup>11</sup> miɿ<sup>25</sup>kai<sup>55</sup>  
deer and monkey or what  
‘A deer and a monkey or something.’ [G1F1-10]
- b. “怎呢 个 涂 合 天 歪..歪 相辅 俺?”  
“tɕɔ<sup>52</sup>ni<sup>55</sup> kai<sup>55-11</sup> t<sup>h</sup>ɔu<sup>55</sup> kaɿ<sup>5</sup> t<sup>h</sup>i<sup>33</sup> bɔi<sup>24-11</sup> ..bɔi<sup>24-11</sup> siɔ<sup>33</sup>hu<sup>24-11</sup> naŋ<sup>52</sup>?”  
why DEF dirt and sky NEG.can..NEG.can help 1PL  
“‘Why can’t the dirt and sky help us?’” [G2F1US-17]
- c. 伊 个 妣 合 伊 个 丈人 合 伊 个 团..  
i<sup>33</sup> kai<sup>55-11</sup> bɔu<sup>52</sup> kaɿ<sup>2-5</sup> i<sup>33</sup> kai<sup>55-11</sup> ti<sup>5</sup>naŋ<sup>55</sup> kaɿ<sup>2-5</sup> i<sup>33</sup> kai<sup>55-11</sup> kiã<sup>52</sup>..  
3SG LW wife and 3SG LW wife's.father and 3SG LW kid..  
睇 ∅ 来 歪.. 清心 □  
thɔi<sup>52</sup> ∅ lai<sup>55</sup> bɔi<sup>24-11</sup> .. tɕ<sup>h</sup>en<sup>33</sup>sim<sup>33</sup> nia<sup>33</sup>  
see 3SG come NEG.can .. content PRT  
‘His wife and his wife’s father and his kid are not happy to see [him] come.’ [G3F2-4]

The conjunction, 合 *kaɿ<sup>2</sup>* ‘and’, is not required in CT Conjunctive Coordination

Constructions of referents. Thus, two or more referents can be coordinated using the asyndetic strategy, as in (60). Note that in (60a), the object 伊 *i<sup>33</sup>* ‘3SG.POSS’ modifies both of the coordinands.

- (60) a. 伊,伊 妈妈 爸爸 画 图 画  
i<sup>33</sup>, i<sup>33</sup> ma<sup>55-11</sup>ma<sup>55</sup> pa<sup>55-11</sup>pa<sup>55</sup> ue<sup>11</sup> tɔu<sup>11</sup> ue<sup>11</sup>  
3SG.POSS mother father paint picture paint  
‘His mother and father are painting a picture.’ [G1F1-10]
- b. 伊 死 凝症 睇见 鸟 獠\*鼠 鱼  
i<sup>33</sup> i<sup>24-11</sup> kuŋ<sup>11</sup>tɕeŋ<sup>21</sup> t<sup>h</sup>ɔi<sup>52-11</sup>kiaŋ<sup>21-53</sup> tɕiau<sup>52</sup> ŋiau<sup>11</sup>tɕ<sup>h</sup>u<sup>52</sup> hu<sup>55</sup>  
3SG sleep dream see bird mouse fish  
‘He sleeps and sees a bird, mouse, and a fish in his dream.’ [G3F1-3]

- c. 只 个 父 只 个 团 坐 食 碱酸  
 tei<sup>52</sup> kai<sup>55-11</sup> pe<sup>24</sup> tei<sup>52</sup> kai<sup>55-11</sup> kiã<sup>52</sup> tɕɔ<sup>24-11</sup> tɕiaŋ<sup>5-2</sup> kiam<sup>11</sup> suŋ<sup>33</sup>  
 this LW father this LW child sit eat food  
 ‘This father and this child sit eating food.’ [G3F1-3]

Based on the evidence in the current data set, 合 *kaŋ<sup>2</sup>* ‘and’ is only used to connect two or more CT ObjRs. It is not used to conjoin two or more clauses. This might not be the case in Singapore Teochew as Low (2014, p. 61) provided the example in (61), where two predicates are coordinated with 合 *kaŋ<sup>2</sup>* ‘and’.

- (61) tiɔŋ<sup>2</sup> ai<sup>53</sup> tak<sup>2</sup> dzik<sup>5</sup> tiaŋ<sup>33</sup> tiaŋ<sup>33</sup> pau<sup>33</sup> kaŋ<sup>2</sup> bua<sup>53</sup> iɔŋ<sup>5</sup> la<sup>53</sup>  
 must every day always wrap and apply medicine PRT  
 iŋ<sup>33</sup> ui<sup>33</sup> tsiŋ<sup>33</sup> tseŋ<sup>53</sup>  
 because very swell  
 ‘He had to wrap his ankle and apply medicine every day because it was very swollen.’ (Low, 2014, p. 61)

There is evidence that Cambodian Teochew borrows the Khmer coordinator ហើយ *haøy* ‘and’, which is used in Khmer Conjunctive Coordinate Clause Constructions to coordinate two or more clauses/events (Sak-Humphry, 2016, p. 230). In Cambodian Teochew, this word is pronounced as *hai<sup>33</sup>* or *hei<sup>33</sup>*, and was shown to be used by three speakers, G1F1, G1F2, and G3F2. The examples in (62) might not be considered complex sentences, per se, because the clauses introduced by ហើយ *hai<sup>33</sup>/hei<sup>33</sup>* seem like they are independent utterances. There is a significant pause between these utterances and the prior utterances. However, each utterance represents an event that is part of a complex whole/figure (Wierzbicka, 1980). The contexts for the examples in (62) was that the respective speakers were talking about one part of the picture,

and then moved to talk about another area with the use of ហើយ *hai<sup>3</sup>/hei<sup>33</sup>* ‘and’. They were thus adding to the complex figure that represents the picture as a whole.

- (62) a. 两 翁妘 在 食 酒 孛团 在 死。  
*nɔ<sup>24-11</sup> aŋ<sup>33</sup>-bɔu<sup>52</sup> tɔ<sup>11</sup> tɕiaɪ<sup>ʔ5-2</sup> tɕiu<sup>52</sup> nɔu<sup>33</sup>kiã<sup>52</sup> tɔ<sup>11</sup> i<sup>24</sup>.*  
 two husband-wife PROG drink alcohol child PROG sleep  
 ហើយ 只个.. 伊伊伊- 媻妘团 在 踏 脚车 去 了  
*hai<sup>33</sup> tɕei<sup>55</sup>.. i<sup>33</sup> i<sup>33</sup> i<sup>33</sup>- tɕa<sup>33</sup>bɔu<sup>52-24</sup>kiã<sup>52</sup> tɔ<sup>11</sup> taɪ<sup>ʔ5-2</sup> k<sup>h</sup>a<sup>33</sup>tɕ<sup>h</sup>ia<sup>33</sup> k<sup>h</sup>u<sup>21</sup> ɔu<sup>11</sup>*  
 and PDCO.. 3SG- girl PROG ride bicycle go already  
 ‘The husband and wife are drinking alcohol and the child is sleeping.  
 And this one, she- [a] girl was riding a bike away.’ [G1F1-10]
- b. 只个 老 人 伊 在 卖 毖毖毖 口。  
*tɕei<sup>55</sup> lau<sup>24-11</sup> naŋ<sup>55</sup> i<sup>33</sup> tɔ<sup>11</sup> bɔi<sup>11</sup> piŋ<sup>33</sup>pɔŋ<sup>33</sup> ɔ<sup>24</sup>.*  
 PDCO old person 3SG PROG sell balloon PRT  
 ហើយ 只个 孛团 伊 读书 出 来 ...  
*hai<sup>33</sup> tɕia<sup>53</sup> nɔu<sup>33</sup>kiã<sup>52</sup> i<sup>33</sup> t<sup>h</sup>ak<sup>5-2</sup>tɕu<sup>33</sup> tɕ<sup>h</sup>uk<sup>2</sup> lai<sup>55</sup> ...*  
 and this.LW child 3SG school exit come ...  
 ‘[In] this (one [the picture]), an old person, he is selling balloons. And this child, she leaves school...’ [G1F1-12]
- c. 伊 又 拍 伊 个 妘 合 伊 个 团  
*i<sup>33</sup> iu<sup>11</sup> p<sup>h</sup>aɪ<sup>ʔ2-5</sup> i<sup>33</sup> kai<sup>55-11</sup> bɔu<sup>52</sup> kaɪ<sup>ʔ2-5</sup> i<sup>33</sup> kai<sup>55-11</sup> kiã<sup>52</sup>*  
 3SG again hit 3SG LW wife with 3SG LW child  
 个 平 伊 食 酒 水。  
*kai<sup>55-11</sup> pɛ<sup>55-11</sup> i<sup>33</sup> tɕiaɪ<sup>ʔ5-2</sup> tɕiu<sup>52</sup> tɕui<sup>52</sup>.*  
 CL same 3SG eat alcohol water  
 ហើយ 有 只 个 老 睇见  
*hei<sup>33</sup> u<sup>11</sup> tɕei<sup>52</sup> kai<sup>55-11</sup> lau<sup>24</sup> thoɪ<sup>52-11</sup>kiaŋ<sup>21</sup>*  
 and have this LW old see  
 ‘He hits his wife with his child again and the same [man], he drinks alcohol. And there is this old [man] who sees/watches.’ [G3F2-4]

- d. 猫.. 坐 口 伊 望 啊.  
 ɲiau<sup>33</sup>.. tɕɔ<sup>24</sup> na<sup>55</sup> i<sup>33</sup> mo<sup>11</sup> a<sup>55</sup>.  
 cat.. sit PRT 3SG gaze PRT.  
 伊 望. 𠵿𠵿 只个 伊 死 啊  
 i<sup>33</sup> mo<sup>11</sup>. hai<sup>33</sup> tɕi<sup>52</sup>kai<sup>55</sup> i<sup>33</sup> uk<sup>5</sup> a<sup>33</sup>  
 3SG gaze. and PDCO 3sg sleep PRT  
 ‘The cat sits and gazes. He gazes. And [in] this (one [the picture])  
 [pointing], he sleeps.’ [G1F2-14]

Speaker G3F2 also uses 𠵿𠵿 *hei<sup>33</sup>* ‘and’ in conjunction with 合 *ka<sup>2</sup>* when coordinating two or more referents, as in (63). According to Huffman (1967, pp. 115, 228), 𠵿𠵿-𠵿𠵿 *haəy-niŋ* ‘with-and’ or just 𠵿𠵿 *niŋ* ‘and’ can be used for Khmer Conjunctive Coordination Constructions with two or more referents. 𠵿𠵿 *haəy* ‘with’ would not be used alone in this construction. Speaker G3F2 appears to have replaced the 𠵿𠵿 *niŋ* ‘and’ part with the Cambodian Teochew word of the same meaning 合 *ka<sup>2</sup>* in (63). However, it is hard to know for sure without more data. It would be interesting to see the extent of the usage of 𠵿𠵿 *hai<sup>3</sup>/hei<sup>33</sup>* ‘and’ amongst other Cambodian Teochew speakers. For example, while speaker G3F2 used this conjunction nearly 25 times, in the same session, speaker G3F3 never used it.

- (63) 只个 马打 伊 互 口.. 只 个 鞋 𠵿𠵿 合 裤  
 tɕia<sup>53</sup> ma<sup>33</sup>ta<sup>24</sup> i<sup>33</sup> hɔu<sup>11</sup> a<sup>33</sup>.. tɕi<sup>52</sup> kai<sup>55-11</sup> ɔi<sup>55</sup> hei<sup>33</sup> ka<sup>2-5</sup> k<sup>h</sup>ɔu<sup>21</sup>  
 this.LW police 3SG give INTJ.. this LW shoe and and pants  
 𠵿𠵿 合 衫 去 只 个 口.. 𠵿𠵿 或 囚犯  
 hei<sup>33</sup> ka<sup>2-5</sup> sã<sup>33</sup> k<sup>h</sup>u<sup>21-53</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> um.. *neak tooh* or prisoner  
 and and shirt go this LW INTJ.. prisoner or prisoner  
 ‘This police officer, he gives these shoes and these pants and this shirt to this  
 prisoner.’ [G3F2-4]

As an alternative to using 侪侪 *hai<sup>3</sup>/hei<sup>33</sup>* ‘and’, CT Conjunctive Coordinate Clause

Constructions are often expressed using the asyndetic strategy, with no overt coordinator, as in (64).

- (64) a. 我我.. 开 门 睇 着 鸡  
 ua<sup>52</sup> ua<sup>52</sup>.. k<sup>h</sup>ui<sup>33</sup> muŋ<sup>55</sup> thoŋ<sup>52</sup> tiɔʔ<sup>2</sup> kɔi<sup>55</sup>  
 1SG open door see RC chicken  
 ‘I opened [the] door and saw a chicken.’ [G2M1-5]
- b. 伊 想 只 鸭 想 獠\*鼠  
 i<sup>33</sup> siɔ<sup>24-11</sup> tɕi<sup>52</sup> aʔ<sup>2</sup> siɔ<sup>24-11</sup> ŋiau<sup>24</sup>tɕ<sup>h</sup>u<sup>52</sup>  
 3SG think this duck think mouse  
 ‘It thinks of this duck and thinks of a mouse.’ [G1F1-9]
- c. Ø 坐 在 空地 在 睇 电视  
 Ø tɕɔ<sup>24-11</sup> tɔ<sup>11</sup> k<sup>h</sup>aŋ<sup>33</sup>-ti<sup>11</sup> tɔ<sup>11</sup> thoŋ<sup>52-24</sup> tian<sup>11</sup>si<sup>24</sup>  
 3SG sit LOC empty-ground PROG watch TV  
 ‘He sits in the clearing and is watching TV.’ [G1F3-15]
- d. 孖团 伊 转 内 伊 合 妈妈 咁  
 nou<sup>11</sup>kiã<sup>52</sup> i<sup>33</sup> tuŋ<sup>52-24</sup>lai<sup>24</sup> i<sup>33</sup> kaʔ<sup>2-5</sup> ma<sup>55-11</sup> ma<sup>55</sup> tã<sup>21</sup>  
 child 3SG return home 3SG with mama talk  
 伊 咁 “妈妈 ...”  
 i<sup>33</sup> tã<sup>21-53</sup> “ma<sup>55-11</sup>ma<sup>55</sup> ...”  
 3SG say mama...  
 ‘The child, he returns home and talks with his mother and says,  
 “Mama...”’ [G2F1US-19]

According to Croft (forthcoming, p. 436), a subtype of conjunctive coordination constructions is a construction that functions to express consecutive temporal events. When CT Conjunctive Coordinate Clause Constructions express a temporal relation, the order of the two clauses is iconic, in that they happen in the order they appear. CT Constructions of this type can use the asyndetic strategy, with no overt marking, as in (59).

- (65) a. 只个 扶\* 簿 来 伊 睇 着 弟弟 在 读  
 tɕi<sup>52</sup>kai<sup>55</sup> iau<sup>11</sup> p<sup>h</sup>ɔu<sup>24</sup> lai<sup>55</sup> i<sup>33</sup> tho<sup>i</sup><sup>52-24</sup> tiɔ<sup>ɔ</sup><sup>5</sup> di<sup>di</sup> tɔ<sup>11</sup> tak<sup>5</sup>  
 PDCO open book come 3SG see RC boy PROG study  
 ‘[In] this (one [the picture]), she opened the book and (then) saw a boy studying.’ [G1F1-12]
- b. □ 伊 欲.. 觅 头家 食 物件  
 ɔ<sup>21</sup> i<sup>33</sup> ai<sup>21</sup>.. tɕ<sup>h</sup>ue<sup>11</sup> t<sup>h</sup>au<sup>55-11</sup>ke<sup>33</sup> tɕia<sup>ɔ</sup><sup>5</sup> mue<sup>ɔ</sup><sup>5-2</sup>kiã<sup>24</sup>  
 INTJ 3SG want.. find owner eat thing  
 ‘It wants to find its owner and (then) eat things.’ [G1F1-9]
- c. 伊 遇 着 伊 个 妘 合 伊 个 团人  
 i<sup>33</sup> ŋɔ<sup>24</sup> tiɔ<sup>ɔ</sup><sup>5</sup> i<sup>33</sup> kai<sup>55-11</sup> bou<sup>52</sup> ka<sup>ɔ</sup><sup>2-5</sup> i<sup>33</sup> kai<sup>55-11</sup> kiã<sup>52-11</sup>-naŋ<sup>55</sup>  
 3SG meet RC 3SG LW wife with 3SG LW child-person  
 伊 在 合 Ø 咁 咁 伊 知神 了  
 i<sup>33</sup> tɔ<sup>11</sup> ka<sup>ɔ</sup><sup>2</sup> Ø tã<sup>21</sup> tã<sup>21-53</sup> i<sup>33</sup> tɕ<sup>h</sup>ai<sup>21-53</sup>seŋ<sup>21</sup> ɔu<sup>11</sup>  
 3SG PROG with 3SG say COMP 3SG realize already  
 ‘He met his wife with his kid and (then) he is telling her that he has realized [his mistakes].’ [G3F2-4]

CT Conjunctive Coordinate Clause Constructions can also use the syndetic strategy with the coordinator 就 *tɕiu*<sup>11</sup> ‘then’. The examples in (66) show 就 *tɕiu*<sup>11</sup> ‘then’ occurring in the break between the two clauses. In (67), 就 *tɕiu*<sup>11</sup> ‘then’ appears after the first argument of the second clause.

- (66) a. 伊 食 好 就 伊 欢喜 死  
 i<sup>33</sup> tɕia<sup>ɔ</sup><sup>5-2</sup> ho<sup>52</sup> tɕiu<sup>11</sup> i<sup>33</sup> huã<sup>33</sup>hi<sup>52</sup> i<sup>24</sup>  
 3SG eat good then 3SG happy sleep  
 ‘He has eaten then he is happy and sleeps.’ [G3F1-3]

- b. 伊 爬 起 来 就 隻 狗..  
 i<sup>33</sup> peŋ<sup>2</sup> k<sup>h</sup>i<sup>52</sup> lai<sup>55</sup> tɕiu<sup>11</sup> tɕiaŋ<sup>2-5</sup> kau<sup>52</sup>..  
 3SG climb up come then CLF.IK dog..  
 被 在 伊 头 顶  
 p<sup>h</sup>ɔŋ<sup>2</sup> tɔ<sup>11</sup> i<sup>33</sup> t<sup>h</sup>au<sup>55-11</sup> teŋ<sup>52</sup>  
 hang LOC 3SG.POSS head atop  
 ‘He gets up and then the dog hangs on top of his head.’ [G1F1-11]
- c. 只个 公 挈 信 好 就 伊 行 口.. 海 舷  
 tɕia<sup>53</sup> kəŋ<sup>33</sup> k<sup>h</sup>iɔŋ<sup>5-2</sup> sunŋ<sup>21</sup> hɔ<sup>52</sup> tɕiu<sup>11</sup> i<sup>33</sup> kia<sup>55-11</sup> na<sup>11</sup>.. hai<sup>52-11</sup> k<sup>h</sup>i<sup>55</sup>  
 this.LW old.man take letter good then 3SG walk LOC.. beach  
 ‘Once this old man takes the letter, then he walks on the beach.’  
 [G3F1-1]
- (67) a. 伊 听 了 伊 就 死 去  
 i<sup>33</sup> t<sup>h</sup>iã<sup>33</sup> liau<sup>52</sup> i<sup>33</sup> tɕiu<sup>11</sup> uk<sup>5</sup> k<sup>h</sup>u<sup>21</sup>  
 3SG listened already 3SG then sleep go  
 ‘He listened then went to sleep.’ [G1F4-16]
- b. 伊.. 伊 死 了 伊 妈妈 就 走 了  
 i<sup>33</sup>.. i<sup>33</sup> i<sup>24</sup> liau<sup>52</sup> i<sup>33</sup> ma<sup>33</sup>ma<sup>33</sup> tɕiu<sup>11</sup> tɕau<sup>52</sup> ɔu<sup>11</sup>  
 3SG 3SG sleep already 3SG.POSS mama then leave already  
 ‘Once he was sleeping, then his mother has left.’ [G1F1-10]
- c. 伊 乏 伊 就 死 了 啊  
 i<sup>33</sup> hek<sup>5</sup> i<sup>33</sup> tɕiu<sup>11</sup> i<sup>24</sup> ɔu<sup>11</sup> a<sup>55</sup>  
 3SG tired 3SG then sleep already PRT  
 ‘He was tired then went to sleep.’ [G1F1-10]

In disjunctive coordination constructions, the coordinands are presented as alternatives (Croft, forthcoming, p. 436). CT Disjunctive Coordination Constructions are expressed using the monosyndetic strategy with the conjunction 啊是 *a<sup>11</sup>si<sup>11</sup>* ‘or’.<sup>90</sup> The examples in (68) show

<sup>90</sup> Mandarin Disjunctive Coordination Constructions are expressed with 还是 *háishi* ‘or’ while Mandarin Alternative Questions are expressed with 或者 *hùozhe* ‘or’ (Li & Thompson, 1981, p. 653). Evidence from Singapore and Jieyang Teochew shows both constructions are co-expressed with 啊是 *a<sup>11</sup>si<sup>11</sup>/a<sup>33</sup>si<sup>35</sup>*

the coordination of referents, and (69) shows an example of a CT Disjunctive Coordinate Clause Construction.

- (68) a. 伊 睇 着 只个.. □ 簿 啊是 乜个  
*i*<sup>33</sup> *thoŋ*<sup>52-24</sup> *tiəŋ*<sup>2</sup> *teŋ*<sup>52</sup>*kai*<sup>55</sup>.. *u*<sup>33</sup> *p<sup>h</sup>əu*<sup>24</sup> *a*<sup>11</sup>*si*<sup>24</sup> *miŋ*<sup>55</sup>*kai*<sup>55</sup>  
 3SG see RC PDCO.. INTJ book or what  
 ‘She saw this one, um a book or something.’ [G1F1-12]
- b. 鹿 合 老猴 啊是 乜个  
*tek*<sup>5</sup> *kaŋ*<sup>2</sup> *lau*<sup>24-11</sup>*kau*<sup>55</sup> *a*<sup>11</sup>*si*<sup>11</sup> *miŋ*<sup>55</sup>*kai*<sup>55</sup>  
 deer and monkey or what  
 ‘A deer and a monkey or something.’ [G1F1-10]
- (69) 食 比萨/ភីហ្សា 啊是 食 乜个 □  
*teŋ*<sup>ŋ</sup><sup>5-2</sup> *p<sup>h</sup>i*<sup>33</sup>*sa*<sup>11</sup> *a*<sup>11</sup>*si*<sup>11</sup> *teŋ*<sup>ŋ</sup><sup>5-2</sup> *mia*<sup>55</sup> *nia*<sup>33</sup>  
 eat pizza or eat something PRT  
 ‘[Someone] is eating pizza or eating something.’ [G1F1-10]

In adversative coordination constructions, the coordinands are construed as being in contrast (Croft, forthcoming, p. 437). There is some evidence that CT Adversative Coordinate Clause Constructions can be expressed with a monosyndetic strategy with 倒 *tə*<sup>53</sup> ‘but’, as in (70) or 但是 *taŋ*<sup>11</sup>*si*<sup>11</sup> ‘but’ in (71). However, given the lack of further examples, more evidence is needed on the usage of these conjunctions. The examples in (70) and (71) are also examples of CT Concessive Constructions, where the events in both clauses happened but the relationship between them is in some way unexpected (Croft, forthcoming, p. 498-9).

---

‘or’ (Low, 2014; Xu, 2007). Data is needed on CT Alternative Questions to see if this is the case for Cambodian Teochew as well.

- (70) 伊人 只 个 组 成立 起 来 有 蜀二年 啊  
 i<sup>33</sup> naŋ<sup>11</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> tɕu<sup>55</sup> seŋ<sup>11</sup> lip<sup>5</sup> ki<sup>52-24</sup> lai<sup>55</sup> u<sup>11</sup> tɕek<sup>2</sup> nɔ<sup>24</sup> ni<sup>55</sup> a<sup>33</sup>  
 3PL this LW group establish up come have one two year PRT  
 倒 无 人 来 唱  
 tɔ<sup>53</sup> bɔ<sup>11</sup> naŋ<sup>55-11</sup> lai<sup>55-11</sup> tɕ<sup>h</sup>iaŋ<sup>21</sup>  
 but NEG people come sing  
 ‘They established this group 1-2 years ago but no one has ever come to sing  
 [Teochew opera]’ [G1F6-24]

- (71) 伊 欢喜 觅 挈 掠 只 獠\*鼠  
 i<sup>33</sup> huã<sup>33</sup> hi<sup>52</sup> tɕ<sup>h</sup>ue<sup>11</sup> k<sup>h</sup>iɔ<sup>ʔ5-2</sup> lia<sup>ʔ5-2</sup> tɕi<sup>52</sup> ŋiau<sup>11</sup> tɕ<sup>h</sup>u<sup>52</sup>  
 3SG happy find take catch this mouse  
 但是 土墙 卡\* 危  
 taŋ<sup>11</sup> si<sup>11</sup> t<sup>h</sup>ɔu<sup>11</sup> tɕ<sup>h</sup>iɔ<sup>55</sup> k<sup>h</sup>a<sup>ʔ2-5</sup> k<sup>h</sup>ui<sup>55</sup>  
 but wall too high  
 就 伊.. 跋 落 来 入.. 𠵼𠵼 底畔  
 tɕiu<sup>11</sup> i<sup>33</sup>.. puã<sup>ʔ5</sup> lɔ<sup>ʔ5</sup> lai<sup>55</sup> dzip<sup>5-2</sup>.. p<sup>h</sup>ɔk<sup>5</sup> tɔi<sup>52-11</sup> pa<sup>ʔ55</sup>  
 then 3SG.. fall descend come enter.. mud inside  
 ‘He is happy to find this mouse to catch, but the wall is too high so he falls  
 down into the mud.’ [G3F1-3]

#### 7.1.5.2 CT Equative and Similitive Constructions

In equative constructions, two unique objects are said to be in the same position on a predicative scale on which they are being compared (Croft, forthcoming, p. 486).

There was one example of a CT Equative Construction that was equating items on the scale of a property, 大 *tua<sup>11</sup>* ‘big’, in (72). Future research should procure more examples of constructions of this type.<sup>91</sup>

<sup>91</sup> But see Xu (2007, pp. 285-7) for Jieyang Equative Constructions.

- (72) 狼 个 平样 大 只 个 狗 平样  
 laŋ<sup>55-11</sup> kai<sup>55-11</sup> pẽ<sup>55-11</sup>iõ<sup>11</sup> tua<sup>11</sup> tçi<sup>52</sup> kai<sup>55-11</sup> kau<sup>52</sup> pẽ<sup>55-11</sup>iõ<sup>11</sup>  
 wolf IK same big this LW dog same  
 ‘[The] wolf is as big as this dog.’ [G2F1US-18]

While equative constructions express equality in quantity, similitive constructions express similarity in quality (Haspelmath with Buchholz, 1998, p. 313). CT Similitive Constructions are given in (73), which show the items being simulated with the predicate 肖 *siau*<sup>24</sup> ‘look like’ using the equal similitive strategy with 平样 *pẽ*<sup>55-11</sup>*iõ*<sup>11</sup> ‘same’ appearing at the end of the utterance.

- (73) a. 圆 肖 日 平样  
 ĩ<sup>55</sup> siau<sup>24-11</sup> dzik<sup>5</sup> pẽ<sup>55-11</sup>iõ<sup>11</sup>  
 circle look.like sun same  
 ‘The circle looks like the sun.’ [G1F1-PS]
- b. 小蜗牛 肖 骨 平样  
 xiǎo-wōniú siau<sup>24-11</sup> kuk<sup>2</sup> pẽ<sup>55-11</sup>iõ<sup>11</sup>  
 small-snail look.like bone same  
 ‘The snail looks like the bone.’ [G2F1US-15]
- c. 汝 肖.. 肖 只 个 花.. 苞 花 平样  
 lu<sup>52</sup> siau<sup>24-11</sup>.. siau<sup>24-11</sup> tçi<sup>52</sup> kai<sup>55-11</sup> hue<sup>33</sup>.. p<sup>h</sup>au<sup>33</sup> hue<sup>33</sup> pẽ<sup>55-11</sup>iõ<sup>11</sup>  
 2SG look.like.. look.like this LW flower.. stalk flower same  
 ‘You look like this flower.’ [G2F1US-19]

### 7.1.5.3 CT Conditional Constructions

Conditional constructions are made up of two clauses, the protasis or antecedent, and the apodosis or consequent (Croft, forthcoming, p. 491). In languages of the world, typically the protasis precedes the apodosis (Croft, forthcoming, p. 496). This is the case for CT Conditional Constructions, as in (74). The example in (74a) uses a syndetic strategy where

morpheme 就 *tɕiu*<sup>11</sup> ‘then’ indicates the conditional relationship between the two clauses.<sup>92</sup>

(74a) exemplifies a content (causal) relation (Croft, forthcoming, p. 492). The construction in (74b) uses an asyndetic strategy, with no additional morpheme identifying the relationship between the clauses. Because of this, the same utterance can have different interpretations in isolation. A hearer will rely on context to determine the appropriate interpretation. (74b) showcases a speech act relation (Croft, forthcoming, p. 492). The speaker’s offer of not collecting a fee is conditional on the potential acceptability of the offer (if there is anyone who wants to come learn) (Sweetser, 1990, pp. 119-20). In both examples in (74), the first clause is a protasis, and the constructions are CT Conditional Constructions based on the contexts in the respective texts.

- (74) a. “汝 摘 伊 伊 就 恁 晓 笑”  
 “lu<sup>52</sup> tia<sup>ʔ2-5</sup> i<sup>33</sup> i<sup>33</sup> tɕiu<sup>11</sup> bɔi<sup>24-11</sup> hiau<sup>52-24</sup> tɕ<sup>hiɔ</sup><sup>21</sup>”  
 2SG pick 3SG 3SG then NEG.can able smile  
 “[If] you pick it, then it won’t be able to smile.” [G2F1US-19]

---

<sup>92</sup> See E. S. Chen (2003) for a similar use of 就 *jiu* ‘then’ in Singapore Colloquial Mandarin Conditional Constructions.

- b. 恁 有 底 啲 欲 来 学习, 欲  
 niŋ<sup>52</sup> u<sup>25-11</sup> ti<sup>11</sup>tiaŋ<sup>55</sup> ai<sup>21-53</sup> lai<sup>55-11</sup> hak<sup>5-2</sup>sip<sup>5</sup>, ai<sup>21-53</sup>  
 2PL have who want come learn want  
 来 练 只 个 潮州 曲 □ ...  
 lai<sup>55-11</sup> liaŋ<sup>11</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> ti<sup>55-11</sup>tɕiu<sup>33</sup> k<sup>h</sup>ek<sup>2</sup> nɔ<sup>33</sup> ...  
 come practice this LW Teochew song PRT ...  
 俺 在 无 价 收  
 naŋ<sup>52</sup> tɔ<sup>11</sup> bɔ<sup>11</sup> ke<sup>21-53</sup> siu<sup>33</sup>  
 1PL PROG NEG fee collect  
 ‘[If] you (pl.) have anyone who wants to come learn, come practice this  
 Teochew song...we are not collecting a fee.’ [G1F6-24]

## 7.2 CT Object Predication Constructions

Stassen (1997, p. 102) uses the file metaphor to distinguish three ways of packaging information of referents: presentational, predicational, and equational. The basis of the metaphor is that our knowledge of the world is made up of different mental files on referents. The presentational information packaging creates a new mental file for a referent. The predicational information packaging adds new information to an already existing file on a referent. The equational information packaging deletes a file because it orders the combination of two unique files into one. We cover predicational constructions in §7.2.1. Croft (forthcoming, p. 267) combines the presentational and equational information packaging functions into the term nonpredicational clause constructions, which are covered in §7.2.2.

### 7.2.1 CT Predicational Constructions

Predicational constructions provide additional information about a referent. In the example of a CT Predicational Construction in (75), the hearer already has a mental file on 巴

士  $pa^{33}sur^{24}$  ‘bus’, and the utterance adds additional information to that file, that the 巴士  $pa^{33}sur^{24}$  ‘bus’ belongs to the 相  $si\sigma^{21}$  ‘picture’ type.<sup>93</sup>

- (75) 只个            巴士            个    相  
 $t\epsilon ia^{53}$              $pa^{33}su^{24}$              $kai^{55}$   $si\sigma^{21}$   
 this.LW            bus            COP    picture  
 ‘This bus is a picture.’ [G3F1-1]

The example in (75) uses the copula strategy, with 个  $kai^{55}$  as a “nonverbal” copula. Stassen (1997) and Croft (forthcoming) differentiate “nonverbal” and “verbal” copula strategies. While historically a generic classifier, we have shown how 个  $kai^{55}$  has been grammaticalized to be used as a linker between EC + referent, deictic attribute + referent, object modifier + referent, action modifier + referent, and even sometimes property modifier + referent. Here we will argue that  $kai^{55}$  has further grammaticalized to act as a linker between

<sup>93</sup> To give context to (75), the mental file for the 巴士  $pa^{33}sur^{24}$  ‘bus’ was established in the picture on the left below, and the picture on the following page shown below on the right provides the additional information that the bus itself is a 相  $si\sigma^{21}$  ‘picture’. The speaker goes on to utter (i).

- (i) 只个            相..            □    电视            底畔  
 $t\epsilon ia^{53}$              $si\sigma^{21}..$              $na^{11}$      $tia\eta^{11}si^{24}$              $t\sigma i^{52-11}pai^{55}$   
 this.LW            picture..            LOC    television            inside  
 ‘This picture is inside a TV.’ [G3F1-1]



referent + object predicate. Per Stassen/Croft’s terminology, the use of 个 *kai*<sup>55</sup> would be considered a “nonverbal” copula strategy.<sup>94</sup> Further examples of CT Predicational Constructions using the “nonverbal” copula strategy are provided in (76).

- (76) a. 伊 个 广告  
*i*<sup>33</sup> *kai*<sup>55</sup> *kuaŋ*<sup>11</sup>*kau*<sup>21</sup>  
 3sg COP advertisement  
 ‘It’s an advertisement.’ [G3F1-1]
- b. 小狐狸 个 狼  
*xiǎohúli* *kai*<sup>55</sup> *lan*<sup>55</sup>  
 fox COP wolf  
 ‘*xiǎohúli* is wolf.’ [G2F1US-18]

In Sinitic languages that lack inflections on predicates, the primary means of distinguishing whether something is “verbal” or not comes from patterns of negation. This is the case with the copula strategies. We currently have no evidence that 个 *kai*<sup>55</sup> as a copula can be negated, hence its qualification as a “nonverbal” copula. On the other hand, there exists another copula in Cambodian Teochew, 是 *si*<sup>24</sup>, that can be negated, as in (77). Thus 是 *si*<sup>24</sup> is considered a “verbal” copula.<sup>95</sup>

- (77) a. 唔 是 唔 是 唔 是 云 口  
*m*<sup>11</sup> *si*<sup>24</sup> *m*<sup>11</sup> *si*<sup>24</sup> *m*<sup>11</sup> *si*<sup>24</sup> *huŋ*<sup>55</sup> *nia*<sup>11</sup>  
 NEG COP NEG COP NEG COP cloud PRT  
 ‘[It] is not a cloud.’ [G1F1-8]

<sup>94</sup> Yap and Matthews (2008, pp. 324-5, citing Matthews and Xu, 2002) similarly state that Teochew 个 *kai* is a “nonverbal” copula in Stassen terms, while 是 *si* (see below) is a “verbal” copula.

<sup>95</sup> See Yue-Hashimoto (1969) for 是 ‘to be’ as a copula in Mandarin.

- b. 唔 是 草 蚱  
 m<sup>11</sup> si<sup>11</sup> tɕ<sup>h</sup>au<sup>52-24</sup> mɛ<sup>52</sup>  
 NEG COP grasshopper  
 ‘[It] is not a grasshopper.’ [G1M1-7]

Other examples of CT Predicational Constructions using the “verbal” copula strategy are given in (78). In both of these examples, the speaker G1M1 does not actually know the information to add to the file.

- (78) a. 唔 知 话 是 乜 个 □  
 m<sup>11</sup> tɕai<sup>33</sup> ue<sup>11</sup> si<sup>11</sup> mi<sup>ʔ5</sup>kai<sup>55</sup> nia<sup>33</sup>  
 NEG know word COP what PRT  
 ‘[I] don’t know what the word is.’ [G1M1-7]
- b. 只 个 是 乜 个  
 tɕi<sup>52</sup>kai<sup>55</sup> si<sup>11</sup> mi<sup>ʔ5</sup>kai<sup>55</sup>  
 PDCO COP what  
 ‘What is this?’ [G1M1-7]

Finally, there appear to be some examples of CT Predicational Constructions where no copula is needed, which Croft (forthcoming, p. 271) calls the zero strategy, as in (79).

- (79) a. 只 隻 鸡 翁 大 隻  
 tɕi<sup>52</sup> tɕia<sup>ʔ2-5</sup> kɔi<sup>33</sup>an<sup>33</sup> tua<sup>11</sup> tɕia<sup>ʔ2</sup>  
 this CLF.ANIMAL rooster big CLF.ANIMAL  
 ‘This rooster is a big one.’ [G1F3-15]
- b. 只 个 乜 个  
 tɕi<sup>52</sup>kai<sup>55</sup> mi<sup>ʔ5</sup>kai<sup>55</sup>  
 PDCO what  
 ‘What is this?’ [G2M1-5, G1M1-7]

### 7.2.2 CT Nonpredicational Clause Constructions

Per Croft (forthcoming, p. 267), presentational and equational information packaging functions are both expressed in nonpredicational clause constructions.

The presentational information packaging is used to introduce a new referent into the discourse context. It is a subtype of thethetic function, which differs from predication in that all of the information is new. In the predicational examples shown above, a referent had already been activated in the discourse context, and information was being added to each referent's file. Conversely, in presentational constructions, a new file for a referent is created.

There was an abundance of CT Presentational Location Constructions in the data set, as in (80). For each of these utterances, the speaker is seeing the referents on the page for the first time and is introducing them, or instructing the hearer to create a new file for each referent. The examples in (80) use the possessive locative strategy (Croft, forthcoming, p. 294), in that they recruit the same form as CT Presentational Possession Construction, which is shown below in (81).

- (80) a. 有 花 有 草 下. 有 兔. 有 狗  
u<sup>11</sup> hue<sup>33</sup> u<sup>11</sup> tɕ<sup>h</sup>au<sup>52</sup> e<sup>24</sup>. u<sup>11</sup> t<sup>h</sup>ɔu<sup>21</sup>. u<sup>11</sup> kau<sup>52</sup>  
have flower have grass below have rabbit have dog  
'There's flowers, there's grass below. There's a rabbit. There's a dog.'  
[G1M1-7]
- b. 有 四 个 人, 四 个 人  
u<sup>11</sup> si<sup>21-53</sup> kai<sup>55-11</sup> naŋ<sup>55</sup>, si<sup>21-53</sup> kai<sup>55-11</sup> naŋ<sup>55</sup>  
have four LW people, four LW people  
'There are four people, four people.' [G3F2-4]

- c. 有 猪 有 马  
u<sup>11</sup> tu<sup>55</sup> u<sup>11</sup> be<sup>52</sup>  
have pig have horse  
‘There is a pig, there is a horse.’ [G2M1-5]
- d. 有 电车. 有 马. 有 鸡 有 猪  
u<sup>11</sup> tian<sup>55-11</sup> tɕia<sup>33</sup>. u<sup>11</sup> be<sup>52</sup>. u<sup>11</sup> kɔi<sup>33</sup> u<sup>11</sup> tu<sup>33</sup>  
have car have horse have chicken have pig  
‘There’s a car. There’s a horse. There’s a chicken, there’s a pig.’  
[G1F1-8]

Presentational possession constructions have three components: the one who possesses something (possessor), the thing possessed (possessum), and the possessive relation between the possessor and possessum (Croft, forthcoming, p. 281). Some CT Presentational Possession Constructions are shown in (81). Stassen (2009) argued that examples like (81) show the use of the topic possessive strategy, in that the possessor, 伊 *i*<sup>33</sup> ‘3SG’ in (81), is the topic and the possessum, 翼 *sik*<sup>5</sup> ‘wing’ in (44a) and 家庭 *ke*<sup>33</sup>*t<sup>h</sup>en*<sup>55</sup> ‘family’ in (81b), is the only argument in the construction. Stassen would gloss a predicate like 有 *u*<sup>11</sup> as ‘exist’. Conversely, Chappell and Creissels (2019) argued that constructions with this structure in Mandarin and other Southeast Asian languages, are using the have possessive strategy, and both the possessor and possessum are arguments of the predicate 有 *u*<sup>11</sup> ‘have’.

- (81) a. 伊 有 翼 ...  
*i*<sup>33</sup> u<sup>11</sup> sik<sup>5</sup> ...  
3SG have wing ...  
‘It [the butterfly] has wings.’ [G2F1US-17]
- b. 伊 有 只 个 好 家庭 □  
*i*<sup>33</sup> u<sup>11</sup> tɕi<sup>52</sup> kai<sup>55-11</sup> hɔ<sup>52-11</sup> ke<sup>33</sup>*t<sup>h</sup>en*<sup>55</sup> e<sup>33</sup>  
3SG have this LW good family PRT  
‘He has this good family.’ [G3F2-4]

In the CT Presentational Location Constructions in (80), the implied location is the page, or the context of the story. There can also be an overt location which is more specific, as in (82).

- (82) a. 水底 有 船  
 tɕui<sup>52-24</sup> tɔi<sup>52</sup> u<sup>11</sup> tɕuŋ<sup>55</sup>  
 water inside have boat  
 ‘Inside the water there is a boat.’ [G1F1-11]
- b. ... 有 鸟 □ 顶 树 顶畔  
 ... u<sup>24</sup> tɕiau<sup>52</sup> na<sup>11</sup> teŋ<sup>52-11</sup> tɕ<sup>h</sup>iu<sup>11</sup> teŋ<sup>52-11</sup> paĩ<sup>55</sup>  
 ... have bird LOC atop tree atop  
 ‘...there is a bird on top of a tree.’ [G3F1-3]
- c. □ 相 底畔 有 马, 有 草  
 na<sup>11</sup> siə<sup>11</sup> tɔi<sup>52-11</sup> pãĩ<sup>55</sup> u<sup>24-11</sup> be<sup>52</sup>, u<sup>24-11</sup> tɕ<sup>h</sup>au<sup>52</sup>  
 LOC picture inside have horse, have grass  
 ‘Inside the picture there is a horse, there is grass.’ [G3F1-1]

CT Presentational Location Constructions also very commonly use a CT Proximal Deictic Contextual ObjR Construction, as in (83). Similar to the utterances in (80), the examples in (83) show the speaker introducing the referents for the first time.

- (83) a. 只个 鸡. 只 鸡 啊  
 tɕei<sup>55</sup> kɔi<sup>33</sup>. tɕei<sup>55</sup> kɔi<sup>33</sup> a<sup>55</sup>  
 PDCO chicken PDCO chicken PRT  
 ‘This (one) [is] a chicken. This (one) [is] a chicken.’ [G1F1-8]

- b. 只 獠\*鼠 啊  
 tɕei<sup>55</sup> ɲiau<sup>24</sup>tɕ<sup>h</sup>u<sup>52</sup> a<sup>33</sup>  
 PDCO mouse PRT  
 ‘This is a mouse.’ [G1F1-9]
- c. 只个 婳妘. 伊 婳妘 问 路 啊  
 tɕei<sup>52</sup>kai<sup>55</sup> tɕa<sup>33</sup>bəu<sup>52</sup>. i<sup>33</sup> tɕa<sup>33</sup>bəu<sup>52</sup> muŋ<sup>11</sup> ləu<sup>11</sup> a<sup>33</sup>  
 PDCO woman IK woman ask road PRT  
 ‘This is a woman. The woman asks for directions.’ [G1M1-7]

In (83), the location is implied, while (84) specifies the location.

- (84) 只个 旅行 船 在 海 中央  
 tɕei<sup>52</sup>kai<sup>55</sup> li<sup>52-24</sup>kiä<sup>55-11</sup> tɕuŋ<sup>55</sup> tɔ<sup>11</sup> hai<sup>52-24</sup> ta<sup>33</sup>uŋ<sup>33</sup>  
 PDCO travel boat LOC sea middle  
 ‘This (one [the picture]) [is] a cruise ship in the middle of the ocean.’ [G1F3-15]

Presentational constructions cannot be negated. The negation of a seemingly presentational turns it into an existential sentence because the referent is now referring to a type, rather than a specific referent (Croft, forthcoming, p. 310). In (85), 船 *tɕuŋ*<sup>55</sup> ‘boat’ is referring to the specific object that the speaker believes is a boat. But once she realizes it is not a boat, 船 *tɕuŋ*<sup>55</sup> ‘boat’ in the following utterance is referring to all boat-type things, and purporting that there are none of those in the picture.

- (85) 水 底 有 船. 无 船 X 涂  
 tɕui<sup>52-24</sup>tɔi<sup>52</sup> u<sup>11</sup> tɕuŋ<sup>55</sup>. bɔ<sup>11</sup> tɕuŋ<sup>55</sup> nia<sup>11</sup> t<sup>h</sup>əu<sup>55</sup>  
 water inside have boat NEG.have boat PRT ground  
 ‘Inside the water there is a boat. There’s not a boat. [It’s] dirt.’ [G1F1-11]

A different type of presentational construction can do more than just present the referent to the hearer. There can also be an additional assertion made about the referent (LaPolla, 1995, pp. 314-316). We will follow Lambrecht (1988, p. 333) to call these CT

Presentational Amalgam Constructions. The examples in (86) all introduce the referent after 有 *u<sup>11</sup>* ‘have’, and then make an additional assertion about said referent.

- (86) a. 有 人 行                      迺迺  
*u<sup>11</sup> naŋ<sup>55</sup> kiã<sup>55-11</sup> t<sup>h</sup>ik<sup>5</sup>t<sup>h</sup>ɔ<sup>55</sup>*  
 have person walk                     leisure  
 ‘There is a person on a stroll.’ [G1F4-16]
- b. 有 飞机                      落                      飞  
*u<sup>11</sup> pue<sup>33</sup>ki<sup>33</sup> lo<sup>ɔ</sup>5-2 pue<sup>33</sup>*  
 have plane descend                     fly  
 ‘There is a plane landing.’ [G1F4-16]
- c. 有 人 数                      耍                      水  
*u<sup>11</sup> naŋ<sup>55</sup> tɕɔi<sup>11</sup> na<sup>11</sup> suŋ<sup>52-11</sup> tɕui<sup>52</sup>*  
 have people many PROG play                     water  
 ‘There are many people playing in the water.’ [G3F1-1]

In (86), the assertion made about the introduced referent was action predication. The examples in (87) show that the assertion in CT Presentational Amalgam Constructions can also be property predication.

- (87) a. 只阵 咋呢 有 老虎                      真真                      大  
*tɕi<sup>52</sup>tɕuŋ<sup>55</sup> tɕɔ<sup>52</sup>ni<sup>55</sup> u<sup>11</sup> lau<sup>24-11</sup>həu<sup>52</sup> tɕiŋ<sup>33</sup>tɕiŋ<sup>33</sup> tua<sup>11</sup>*  
 now why have tiger very RED                     big  
 ‘Now why is there a very big tiger?’ [G2F1US-18]
- b.                      路 有 车                      数 人                      数  
*na<sup>11</sup> lo<sup>u</sup><sup>11</sup> u<sup>24-11</sup> tɕ<sup>h</sup>ia<sup>33</sup> tɕɔi<sup>11</sup> naŋ<sup>55</sup> tɕɔi<sup>11</sup>*  
 LOC road have car many                     people many  
 ‘On the road there are many cars and many people.’ [G3F1-1]

With the equational information packaging function, a hearer already has a file for both things, and an utterance serves to merge the two existing files. There was only one example of

this type of construction, in (88), which uses a nonverbal copula strategy. The example in (88) was uttered after we have already been introduced to the 人 *naŋ*<sup>55</sup> ‘person’ and the 公 *kəŋ*<sup>33</sup> ‘old.man’. We had two separate files for the two referents, and the utterance below is telling the hearer to merge them. More data is needed on the usage of this type of construction in Cambodian Teochew.

- (88) 人    □                    信物                    个    只    个    公  
*naŋ*<sup>55</sup>   *nim*<sup>24-11</sup>                    *suŋ*<sup>21</sup>-*mueŋ*<sup>5</sup>                    *kai*<sup>55</sup>   *tɕi*<sup>52</sup>   *kai*<sup>55-11</sup>   *kəŋ*<sup>33</sup>  
 person hold                    letter-thing                    COP    this    LW    old.man  
 ‘The person holding the letter is this old man.’ [G3F1-1]

### 7.3 CT Property Predication Constructions

Properties can act as predicates. Unlike English, CT Property Predication Constructions do not use the copula strategy. Instead, CT Property Predication Constructions use the “verbal” strategy (Croft, forthcoming, p. 269, citing Stassen 1997). They recruit the same strategy that is used for action predication, and thus is an example of the actual information packaging (IP) strategy. Because of this similarity in the form of action predication and property predication, some have said that a language like Teochew does not have a class of “adjectives” because “adjectives” act as “stative verbs” (Low, 2014, p. 29; Yeo, 2011, p. 33). However, we showed in Chapter 6 that Cambodian Teochew does indeed have prototypical properties that act as modifiers. It is only when properties act as predicates that they resemble action predication. The RCG analysis allows us to easily explain these behaviors.

Examples of CT Property Predication Constructions using the actual IP strategy are shown in (89). There is no use of a copula and there are no auxiliaries. The word order is referent + property, parallel to CT Action Predication Constructions with the order referent + action. It has been argued that for Sinitic languages such as Mandarin (Sun, 2022, pp. 208-9)

and Cantonese (Francis & Matthews, 2005, p. 277), predicated properties without any additional modification such as an intensifier often gives the utterances a feeling of “incompleteness”. Xu (2007, p. 52) similarly argued that in JT Property Predication Constructions, the property must either be reduplicated or contain an additional modifier such as 酷 *hoʔ²* ‘very’. If the same were true for Cambodian Teochew, the use of the resumptive 伊 *i³³* ‘3SG’ in (89a-b) below could be argued to rectify the sentence’s feeling of incompleteness. However, we still see examples such as (89c-e) which solely show the referent and the predicated property. The allowance of unmodified predicated properties in Cambodian Teochew could be another area for contact-induced language change, as properties in Khmer Property Predication Constructions regularly arise without any additional modifiers (Sak-Humphry, 2016, p. 214).

- (89) a. 海船            伊    大  
           *hai*<sup>52-24</sup> *tɕuŋ*<sup>55</sup>    *i*<sup>33</sup>    *tua*<sup>11</sup>  
           sea-boat        3SG    big  
           ‘The boat, it’s big.’ [G1F3-15]
- b. 只    个    口    狐狸,            伊    聪明            啊  
       *tɕi*<sup>52</sup>    *kai*<sup>55-11</sup>    *a*<sup>33</sup>    *hu*<sup>33</sup>*li*<sup>55</sup>,        *i*<sup>33</sup>    *tɕ*<sup>h</sup>*ɔŋ*<sup>33</sup>*mɛŋ*<sup>55</sup>    *a*<sup>33</sup>  
       this    LW    INTJ    fox,            3SG    smart            PRT  
       ‘This fox, he is smart.’ [G2F1US-18]
- c. 只    个    色    雅  
       *tɕi*<sup>52</sup>    *kai*<sup>55-11</sup>    *sek*<sup>5</sup>    *ŋia*<sup>52</sup>  
       this    LW    color    pretty  
       ‘This color is pretty.’ [G1F3-15]

- d. 老猴            𠵹      啊  
 lau<sup>24-11</sup> kau<sup>55</sup>    tɕɔi<sup>11</sup>    a<sup>11</sup>  
 monkey        many    PRT  
 ‘[There are] many monkeys.’ [G1F1-11]
- e. 天顶            月娘            光            啊  
 tʰɿ<sup>33</sup> teŋ<sup>52</sup>      gueŋ<sup>5-2</sup> niŋ<sup>55</sup>    kuŋ<sup>33</sup>        a<sup>11</sup>  
 sky              moon        bright        PRT  
 ‘The sky and the moon are bright.’ [G1F4-16]
- f. 只    隻            新    只    隻            故  
 tɕi<sup>52</sup>    tɕiaŋ<sup>2-5</sup>      siŋ<sup>33</sup>    tɕi<sup>52</sup>    tɕiaŋ<sup>2-5</sup>      ku<sup>11</sup>  
 this    CLF.LARGE    new    this    CLF.LARGE    old  
 ‘This [car] is new, this [car] is old.’ [G1F4-PS]

CT Property Predication Constructions can be modified with CT Intensifier

Constructions using a degree modifier strategy. CT Intensifier 真 *tɕiŋ*<sup>33</sup> ‘very’ is used before the property, as in (90).

- (90) a. 伊    真    雅  
 i<sup>33</sup>    tɕiŋ<sup>33</sup>    ŋia<sup>52</sup>  
 3SG    very    beautiful  
 ‘It is very beautiful.’ [G2F1US-20]
- b. “个    铁    门    真    大.    好    俺    细”  
 “kai<sup>55-11</sup> tiŋ<sup>2-5</sup>    muŋ<sup>55</sup> tɕiŋ<sup>33</sup>    tua<sup>11</sup>.    ho<sup>52</sup>    naŋ<sup>52</sup>    soi<sup>21</sup>”  
 “IK    iron    door    very    big.    good    1PL    small”  
 ‘The iron door is very big and we are small.’ [G2F1US-20]

CT Intensifiers 死 *sɿ*<sup>52</sup> ‘die’ and 绝 *tɕɔŋ*<sup>5</sup> ‘very’ are used after the property, as in (91) and (92).

- (91) a. Ø 欢喜 死  
 Ø huā<sup>33</sup>hi<sup>52-24</sup> si<sup>52</sup>  
 1SG happy die  
 ‘[I’m] very happy.’ [G2M1-5]
- b. 我 疲 死 了, Ø 欲 转 内 了  
 ua<sup>52</sup> hek<sup>5-2</sup> si<sup>52</sup> ou<sup>11</sup>, Ø ai<sup>21-53</sup> tuŋ<sup>52</sup> lai<sup>24</sup> ou<sup>11</sup>  
 1SG tired die already, 1SG want return home already  
 ‘[I am] so tired already. I want to go home.’ [G2F1US-20]
- (92) a. 花 数 绝. Ø 儒 绝  
 hue<sup>33</sup> tɕai<sup>11</sup> tɕɔŋ<sup>5</sup>. Ø dzu<sup>55-11</sup> tɕɔŋ<sup>5</sup>  
 flower many very. 3PL exquisite very  
 ‘There are a lot of flowers. [They are] very exquisite.’ [G1F4-16]
- b. Ø 光 绝  
 Ø kuŋ<sup>33</sup> tɕɔŋ<sup>5</sup>  
 3SG bright very  
 ‘[The fire is] very bright.’ [G1F4-16]

Stassen (1997, p. 45) used negation as a criterion for whether a language uses a “verbal” strategy for non-prototypical predication. Thus if the property predication is negated in the same way action predicates are negated, it is an example of the actual IP strategy being used. In §7.1, we showed strategies for negation of declaratives and the use of several negators in Cambodian Teochew. Some of those same negators are used in CT Property Negation Constructions.

When negating gradient CT Properties, 𠵿 *bɔi*<sup>11</sup> NEG.can is used, as in (93).

- (93) a. 俺 个 身 ... 伊 𠵿 硬 啊  
 naŋ<sup>52</sup> kai<sup>55-11</sup> siŋ<sup>33</sup> ... i<sup>33</sup> bɔi<sup>24-11</sup> ŋɛ<sup>24</sup> a<sup>33</sup>  
 1PL LW body ... 3SG NEG.can hard PRT  
 ‘Our body...it’s not hard.’ [G2F1US-17]

- b. Ø 爬 套 猛  
 Ø pe<sup>55</sup> bɔi<sup>24-11</sup> mẽ<sup>52</sup>  
 1PL crawl NEG.can fast  
 ‘[We] don’t crawl fast.’ [G2F1US-17]
- c. 伊 个 妘 合 伊 个 丈人 合  
 i<sup>33</sup> kai<sup>55-11</sup> bɔu<sup>52</sup> kaɿ<sup>2-5</sup> i<sup>33</sup> kai<sup>55-11</sup> tiʃ<sup>11</sup> naŋ<sup>55</sup> kaɿ<sup>2-5</sup>  
 3SG LW wife with 3SG LW wife's.father with  
 伊 个 团.. 睇 Ø 来 套.. 清心 □  
 i<sup>33</sup> kai<sup>55-11</sup> kiã<sup>52</sup>.. thɔi<sup>52</sup> Ø lai<sup>55</sup> bɔi<sup>24-11</sup>.. tɛ<sup>h</sup>ɛŋ<sup>33</sup>sim<sup>33</sup> nia<sup>33</sup>  
 3SG LW kid.. see 3SG come not.. content PRT  
 ‘His wife and his wife’s father and his kid are not happy to see [him] come.’ [G3F2-4]
- d. 蜀 双 雅 蜀 双 套 雅  
 tɕek<sup>5-2</sup> saŋ<sup>33</sup> ŋia<sup>52</sup> tɕek<sup>5-2</sup> saŋ<sup>33</sup> bɔi<sup>24-11</sup> ŋia<sup>52</sup>  
 one CLF.PAIR pretty one CLF.PAIR NEG.can pretty  
 ‘one pair [of shoes] is pretty, one pair is not pretty’ [G1F4-PS]

According to Xu (2005, 2007), in Jieyang Teochew Property Negation Constructions, 唔 *m*<sup>35</sup> ‘not’ is used to negate non-gradient Jieyang Properties, as in (94).

- (94) a. m<sup>35-21</sup> koŋ<sup>33</sup> peŋ<sup>55</sup> kai<sup>35-21</sup> su<sup>11</sup> kau<sup>213-53</sup> ts<sup>h</sup>u<sup>213</sup> to<sup>33</sup> u<sup>35</sup>  
not fair LW matter everywhere all exit  
 ‘There are unfair phenomena everywhere.’ (Xu, 2005, p. 184)
- b. heŋ<sup>11</sup> tsai<sup>35</sup> nam<sup>55-11</sup> nuŋ<sup>53</sup> p<sup>h</sup>ɛŋ<sup>53</sup> m<sup>35-21</sup> tioɿ<sup>5</sup> tsu<sup>35-21</sup> pai<sup>53-35</sup>  
 now man woman equal not right them put  
 lai<sup>55-11</sup> ts<sup>h</sup>uŋ<sup>55-11</sup>-teŋ<sup>53</sup> p<sup>h</sup>ueɿ<sup>5-2</sup> kau<sup>213-53</sup> tek<sup>5</sup>  
 come table-top talk till straight  
 ‘Nowadays, men and women are equal. If there is any dispute, then they should put it on the table and solve it.’ (Xu, 2007, pp. 221-2)

There was one example of a CT Property Negation Construction with 唔 *m*<sup>11</sup> ‘NEG’ in (95). However, (96) shows the same property from the Jieyang example in (94b) negated with 无 *bɔ*<sup>55</sup> ‘NEG’ in Cambodian Teochew. This could be another instance of preference for the use

of 无 *bɔ*<sup>55</sup> ‘NEG’ as a negator, as was shown in CT Declarative Negation Constructions, however more data is needed to determine the constraints on the use of negators in CT Property Negation Constructions.

- (95) 只个 苹果 伊 平样. 只个 唔 平样  
*tɕi*<sup>52</sup>*kai*<sup>55</sup> *p<sup>h</sup>eŋ*<sup>33</sup>*ku*<sup>52</sup> *i*<sup>33</sup> *pɛ*<sup>55-11</sup>*i*<sup>11</sup>. *tɕi*<sup>52</sup>*kai*<sup>55</sup> *m*<sup>11</sup> *pɛ*<sup>55-11</sup>*i*<sup>11</sup>.  
 PDCO apple 3PL same PDCO NEG same  
 ‘These are apples, they are the same. These are not the same.’ [G1F3-PS]

- (96) Ø 再 想 但.. 行 只 个 路 着,  
 Ø *tɕai*<sup>21-53</sup> *si*<sup>24-11</sup> *tã*<sup>21-53</sup> .. *kiã*<sup>55-11</sup> *tɕi*<sup>52</sup> *kai*<sup>55-11</sup> *lou*<sup>11</sup> *tio*<sup>ʔ5</sup>,  
 3SG again think COMP.. walk this LW road right,  
 只 个 路 无 着  
*tɕi*<sup>52</sup> *kai*<sup>55-11</sup> *lou*<sup>11</sup> *bɔ*<sup>55-11</sup> *tio*<sup>ʔ5</sup>  
 this LW road NEG right  
 ‘He thinks again that taking this path is right, and this path is not right.’  
 [G3F3-4]

We have shown that CT Object Modification uses a copula strategy while both CT Property Modification and CT Action Modification do not use a copula strategy. This conforms

to the universal from Croft (forthcoming pp. 277, citing Croft 1991, p. 130; Stassen 1997, p. 127) in (97).

- (97) a. “If a language has a zero-coded predication construction for object concepts, then it has a zero-coded predication construction for property concepts;  
b. If a language has a zero-coded predication construction for property concepts, then it has a zero-coded predication construction for action concepts.” (Croft, forthcoming, p. 277)

The pattern we see in CT Property Predication also conforms to the Tensedness Universal in Stassen (1997, p. 357, as cited in Croft, forthcoming, p. 277), that says that a non-tensed language will have “verby” property predication. As Cambodian Teochew is a non-tensed language, we would expect it to use the actual IP strategy (“verby”) for property predication and this is the case.

#### 7.4 CT Locative Predication Constructions

Another type of non-prototypical predication is location predication, which is a clause with a location relation. There are three semantic components in locative predication constructions: the referent of interest with the predicated location (the figure); the referent that helps define the locative relation (the ground); and the actual locative relation (the path) (Croft, forthcoming, p. 279). The location of the figure is being predicated in locative predication constructions, where the location is made up of the path and the ground (Croft, forthcoming, p. 279).

Simple CT Locative Predication Constructions consist of a locative predicate, 在 *toʰ* ‘be.at’, without an additional path word, as in (98). These examples are all showing predication

in that each referent, the figure, has already been introduced, and this additional information of the referent's location is being added to its file.

- (98) a. 市场.. 在 电,电车 地方  
 tɕ<sup>hi</sup>i<sup>24-11</sup>ti<sup>55</sup>.. tɔ<sup>11</sup> tian<sup>11</sup>, tian<sup>11</sup>tɕ<sup>hi</sup>ia<sup>33</sup> ti<sup>11</sup>huŋ<sup>33</sup>  
 market be.at car place  
 ‘[The] market is on the parking lot.’ [G1F1-8]
- b. 伊 在 美国 啊  
 i<sup>33</sup> tɔ<sup>11</sup> mui<sup>24</sup>kək<sup>2</sup> a<sup>11</sup>  
 3SG be.at USA PRT  
 ‘He’s in the United States.’ [G2M1-5]
- c. 唔 知 伊 媛 在 块 口  
 m<sup>11</sup> tɕai<sup>33</sup> i<sup>33</sup> ai<sup>55</sup> tɔ<sup>11</sup> kɔ<sup>21</sup> nia<sup>11</sup>  
 NEG know 3SG.POSS mother be.at where PRT  
 ‘I don’t know where his mother is.’ [G2M1-6]
- d. 只 个 丈夫 伊 在..在.. 警察厝  
 tɕei<sup>52</sup> kai<sup>55-11</sup> ta<sup>33</sup>pou<sup>33</sup> i<sup>33</sup> tɔ<sup>11</sup>..tɔ<sup>11</sup>.. jǐngchá-tɕ<sup>hu</sup>u<sup>21</sup>  
 this LW man 3SG be.at..be.at.. police-building  
 ‘This man, he is in prison.’ [G3F3-4]

Much more frequently, CT Locative Predication Constructions are made up of a locative predicate, 在 tɔ<sup>11</sup> ‘be.at’ or 口 na<sup>11</sup> ‘be.at’, with an additional path word, as in (99).

- (99) a. 口 伊 鸟 在 伊 厝 顶  
 ɔ<sup>11</sup> i<sup>33</sup> tɕiau<sup>52</sup> tɔ<sup>11</sup> i<sup>33</sup> tɕ<sup>hu</sup>u<sup>21-53</sup> teŋ<sup>52</sup>  
 INTJ IK bird be.at 3SG.POSS house atop  
 ‘The bird is on top of his house.’ [G1F1-10]
- b. 位洒水物 口 船 顶畔  
 ui<sup>11</sup>-siu<sup>11</sup>tɕui<sup>52</sup>-mue<sup>75</sup> na<sup>11</sup> tɕuŋ<sup>55</sup> teŋ<sup>52-11</sup>paĩ<sup>55</sup>  
 place-swim-thing be.at boat atop  
 ‘[The] swimming pool thing is on top of a boat.’ [G3F1-1]

- c.  只个  船  □  海  顶畔  
 tɕia<sup>53</sup>  tɕuŋ<sup>55</sup>  na<sup>11</sup>  hai<sup>52</sup>  teŋ<sup>52-11</sup>paĩ<sup>55</sup>  
 this.LWboat  be.at  sea  atop  
 ‘This boat is on the sea.’ [G3F1-1]
- d.  大  船  在  海  顶畔  啊  
 tua<sup>11</sup>  tɕuŋ<sup>55</sup>  tɔ<sup>11</sup>  hai<sup>52</sup>  teŋ<sup>52-24</sup>paĩ<sup>55-11</sup>  a<sup>33</sup>  
 big  boat  be.at  ocean  atop  PRT  
 ‘[The] big boat is on top of the ocean.’ [G1F1-8]
- e.  只个  船物  □..  巴士  顶畔  
 tɕia<sup>53</sup>  tɕuŋ<sup>55</sup>-mueŋ<sup>5</sup>  na<sup>11</sup>..  pa<sup>33</sup>su<sup>24</sup>  teŋ<sup>52-11</sup>paĩ<sup>55</sup>  
 this.LWboat-thing  be.at..  bus  atop  
 ‘This boat thing is on a bus.’ [G3F1-1]

Traditionally, in Teochew Locative Predication Constructions, the additional path word, such as 顶畔 *teŋ<sup>52-11</sup>paĩ<sup>55</sup>* ‘atop’, if present, occurs after the ground, as the examples in (99) show. These path words are shown to appear after the ground in Singapore Teochew in Low (2014) and Yeo (2011, p. 46), and in Jieyang (Xu, 2007).

However, McFarland (2017, p. 16) first noted the existence of examples with the path occurring before the ground, adjacent to the locative predicate, as in (100).

- (100) a.  鸟  在  头前  个  大夫团  □  盒  
 tɕiau<sup>52</sup>  tɔ<sup>11</sup>  t<sup>h</sup>au<sup>55-11</sup>tɕai<sup>55</sup>  kai<sup>55-11</sup>  ta<sup>33</sup>pou<sup>33</sup>-kiã<sup>52</sup>  nim<sup>24-11</sup>  ap<sup>5</sup>  
 bird  be.at  front  IK  man-child  hold  box  
 ‘[The] bird is in front of the boy holding a box.’ [G2F1US-PS]  
 (McFarland, 2017, p. 26)
- b.  老姐  耍内物  □  底畔  报纸  
 lau<sup>24-11</sup>-tɕe<sup>52</sup>  suŋ<sup>52-11</sup>lai<sup>24-11</sup>-mueŋ<sup>5</sup>  na<sup>11</sup>  tɔi<sup>11</sup>paĩ<sup>55-11</sup>  pɔ<sup>53</sup>tɕua<sup>52</sup>  
 old-girl  play-house-thing  be.at  inside  newspaper  
 ‘The girl playing house is inside a newspaper.’ [G3F1-1]

In (101), the path is expressed twice, before and after the ground.

- (101) ... 有 鸟 口 顶 树 顶畔  
 ... u<sup>24</sup> tɕiau<sup>52</sup> na<sup>11</sup> ten<sup>52-11</sup> tɕ<sup>h</sup>iu<sup>11</sup> ten<sup>52-11</sup>paĩ<sup>55</sup>  
 ... have bird be.at atop tree atop  
 ‘...there is a bird on top of a tree.’ [G3F1-3]

In §7.4.1 we showed the use of optional path words occurring after the ground in motion events in CT Eventive Complex Predicate Constructions. There is some evidence that path words in these constructions can also be expressed before the ground, as in (102).

- (102) 伊 过 溪 好 就 伊 爬\* 起 顶 树  
 i<sup>33</sup> kue<sup>21-52</sup> k<sup>h</sup>ɔi<sup>33</sup> hɔ<sup>52</sup> tɕiu<sup>11</sup> i<sup>33</sup> peɿ<sup>5</sup> k<sup>h</sup>i<sup>52-11</sup> ten<sup>52-11</sup> tɕ<sup>h</sup>iu<sup>11</sup>  
 3SG cross river good then 3SG climb up atop tree  
 ‘Once he has crossed the river, he climbs up a tree.’ [G3F1-3]

Based on the albeit limited evidence here, a shift might be happening for some speakers where the path word is allowed to be expressed before the ground in some constructions such as CT Locative Predication Constructions and CT Eventive Complex Predicate Constructions. This may be a result of influence from Khmer. In Khmer Locative Predication Constructions, the canonical word order is path + ground, as in (103).

- (103) a. ក្រោម ផ្ទះ មាន គោ  
 kroam pteah mien koo  
 under house have cow  
 ‘Under the house there are cows.’ (Sak-Humphry, 2016, p. 116)
- b. ក្នុង ធុង មាន ទឹក ។  
 knoŋ thuŋ mien tik  
 in container have water  
 ‘In the container there is water.’ (Sak-Humphry, 2016, p. 116)

## 7.5 Summary

This chapter has given a snapshot of some of the most frequent action, object, and property predication constructions in the current data set on Cambodian Teochew. As for CT Action Predication Constructions, we elected not to use the terms “subject” and “direct object” in describing any of the constructions because there was no data to support the existence of such categories in Cambodian Teochew. Most of the constructions focused on declarative speech acts, including the variety of negative morphemes used in CT Declarative Negation Constructions. Though more research is needed in this area, initial evidence points to the increasing use of 𠵼 *bɔ̃<sup>55</sup>* ‘NEG’ in these constructions, as compared to the usage in other Teochew varieties. There were a few phenomena that were hypothesized to have arisen due to contact with Khmer. Limited evidence pointed to the use of lexical 𠵼 *hɔu<sup>11</sup>* ‘give’ in CT Causative Constructions to make an intransitive predicate transitive. Lexical ‘give’ in Khmer is similarly used in Khmer Causative Constructions. We found evidence for a non-canonical word order in regards to the path word and the ground that was not found in other Teochew varieties but matched the order found in Khmer Locative Predication Constructions. Additionally, we found the use of Khmer loanwords 𠵼 *hai<sup>33</sup> / he<sup>33</sup>* ‘and’ and 𠵼 *pi<sup>33</sup>* ‘from’ and showcased their usage in the relevant CT Constructions.

The types of constructions performing the information packaging function of predication are immense in any language. Since a limited number were covered here, there are many areas for future research.

## Chapter 8 – Khmer Loanwords<sup>96</sup>

In §2.6, it was shown that Sinitic languages have influenced the local lexicons of Southeast Asian languages. At the same time, the Chinese diaspora have adopted many loanwords into their Sinitic varieties from the dominant languages of the countries in which they settled. This chapter gives an overview of some of the Khmer loanwords used in Cambodian Teochew by this study's speaker consultants. §8.1 presents the methodology for determining what constituted a loanword, while §8.2 contains the list of loanwords, then §8.3 details how the words are integrated into the Cambodian Teochew language, and §8.4 offers some points of discussion.

### 8.1 Methodology

This portion of the study uses the data detailed in §1.4 from the nine primary speaker consultants and two of the secondary speaker consultants (G2F1US and G2M2US). The assembled data constituted a preliminary vocabulary of Cambodian Teochew which was then compared to sources on other Teochew varieties including Indonesian (IT) (X. Chen, 2014; Peng, 2012), Jieyang (JT) (Xu, 2007), Lao (LT) (X. Chen, 2014), Johor Bahru Teochew of Malaysia (MT) (X. Chen, 2014), Singapore (ST) (X. Chen, 2014; Goh, 2017; Li, 1991; Low, 2014; Yeo, 2011), Thai (TT) (Atchariyasucha, 1982, X. Chen, 2014), and Vietnam (VT) (X. Chen, 2014). 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, 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 CT word did not resemble the form from any of the other

---

<sup>96</sup> This chapter has been adapted from McFarland (2021).

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 §2.6, a final check was made to Pou and Jenner (1973) to confirm the direction of the borrowing. For example, the CT word 𑜁𑜂𑜃𑜂𑜃𑜂𑜃 *tɕ<sup>h</sup>a<sup>24</sup>tɕ<sup>h</sup>au<sup>52</sup>* ‘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, p. 47). Therefore, if a word was found in Pou and Jenner (1973), it was decided that it was not a case of Cambodian Teochew borrowing from Khmer.

## 8.2 Loanwords

After going through the aforementioned process with the collected vocabulary items, it was determined that the items listed in Table 8.1 constitute words 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 under the IPA if they were present in the original source.

Cambodian Teochew	English Gloss	Khmer Etymon <sup>97</sup>	Other Teochew Sources
bɔk <sup>2</sup> lɔ <sup>11</sup> hɔŋ <sup>55</sup>	papaya salad	ប៉ាកល្បង ប៉ាកល្បង	
pɔ <sup>33</sup> la <sup>33</sup> hɔk <sup>5</sup>	prahok (fermented fish paste)	ប្រហុក ប្រហុក	

<sup>97</sup> Filippi and Hiep (2016) and Sak-Humphry (2016) were referenced for Khmer pronunciation.

bəŋ <sup>33</sup> tɕ <sup>h</sup> au <sup>52</sup>	Cambodian crepe, pancake	បាញ់ឆែវ បាញ់ឆែវ	
lɔk <sup>2</sup> laŋ <sup>5</sup> , tɕ <sup>h</sup> a <sup>52-24</sup> lɔk <sup>2</sup> laŋ <sup>5</sup>	lok lak (stir-fried beef cubes)	លក់លាវ ឡុកឡាក់	
a <sup>11</sup> mək <sup>5</sup>	amok (curry in banana leaf)	a:mək អាម៉ុក	
ka <sup>33</sup> laŋ <sup>33</sup>	kralan (bamboo sticky rice)	krɔla:n ក្រឡាន	
maŋ <sup>11</sup> k <sup>h</sup> u <sup>2</sup>	mangosteen, manggis	mɔvəŋk <sup>h</sup> ut មង្គុត	LT: maŋ <sup>33</sup> k <sup>h</sup> uk <sup>5</sup> (X. Chen, 2014, p. 910) MT: maŋ <sup>33</sup> hik <sup>5</sup> (X. Chen, 2014, p. 827) ST: maŋ <sup>55</sup> hek <sup>1</sup> (Goh, 2017, p. 209) 芒黑 maŋ <sup>11-12</sup> hi <sup>2</sup> (X. Chen, 2014, p. 851) TT: maŋ <sup>33</sup> k <sup>h</sup> uk <sup>5</sup> (X. Chen, 2014, p. 922) VT: maŋ <sup>33</sup> ku <sup>2</sup> (X. Chen, 2014, p. 903)
sau <sup>33</sup> mau <sup>33</sup>	rambutan	sa:vma:v សាវម៉ាវ	IT: lam <sup>33</sup> pu <sup>2</sup> taŋ <sup>33</sup> (X. Chen, 2014, p. 864) MT: an <sup>55-33</sup> mo <sup>55-33</sup> taŋ <sup>33</sup> (X. Chen, 2014, p. 827) 红毛丹 ST: aŋ <sup>55-12</sup> mo <sup>55-12</sup> taŋ <sup>33</sup> (Li, 1991, p. 59) 红毛丹 aŋ <sup>55</sup> mo <sup>55</sup> taŋ <sup>33</sup> (Goh, 2017, p. 205) 红毛丹 aŋ <sup>55-33</sup> mo <sup>55-33</sup> tũa <sup>33</sup> (X. Chen, 2014, p. 851) 红毛丹 TT: aŋ <sup>55-33</sup> mo <sup>55-33</sup> tũa <sup>33</sup> (X. Chen, 2014, p. 922) 红毛丹 VT: aŋ <sup>55-33</sup> mo <sup>55-33</sup> tũa <sup>33</sup> (X. Chen, 2014, p. 908) 红毛丹
buu <sup>33</sup>	avocado	p <sup>h</sup> laɛbɔ̃ ផ្លែប៊ីវ	LT: ma <sup>2</sup> sə <sup>33</sup> uə <sup>33</sup> (X. Chen, 2014, p. 913) ŋak <sup>33</sup> lai <sup>55</sup> / gu <sup>11</sup> ni <sup>11</sup> kuɛ <sup>53</sup> (TCKnow, 2015) 鳄梨 / 牛奶果

ka <sup>11</sup> lem <sup>33</sup>	ice cream	ka:reem កាណែម	MT: ai <sup>33</sup> si <sup>33</sup> k <sup>h</sup> i <sup>2</sup> lim <sup>33</sup> (X. Chen, 2014, p. 840)
			ST: ai <sup>33</sup> sə <sup>33</sup> kik <sup>1</sup> lim <sup>55</sup> (Goh, 2017, p. 205) 爱士吉林 ai <sup>33</sup> si <sup>11</sup> k <sup>h</sup> ə <sup>11</sup> lim <sup>33</sup> (X. Chen, 2014, p. 856)
			VT: ka <sup>33</sup> liam <sup>33</sup> (X. Chen, 2014, p. 905)
			suŋ <sup>33</sup> kə <sup>33</sup> / səŋ <sup>55</sup> kə <sup>33</sup> (TCKnow, 2015) 霜膏 / 雪糕
lə <sup>11</sup> həŋ <sup>55</sup>	papaya	lhəŋ លួង	HT: bak <sup>5</sup> kue <sup>33</sup> (Fielde, 1883, p. 277) 木瓜
			IT: pakue / nikue (Peng, 2012, p. 33)
			ST: bak <sup>5</sup> kue <sup>33</sup> / ni <sup>33</sup> kue <sup>33</sup> (Goh, 2017, p. 134) 木瓜 / 奶瓜
pə <sup>33</sup> loi <sup>52</sup>	jungle, forest	prei ព្រៃ	HT: tʃ <sup>h</sup> i <sup>313</sup> -nã <sup>55</sup> -ta <sup>313</sup> / tʃ <sup>h</sup> i <sup>313</sup> -p <sup>h</sup> e <sup>313</sup> / tʃ <sup>h</sup> iu <sup>11</sup> -bak <sup>2</sup> im <sup>33</sup> -ue <sup>313</sup> (Duffus, 1883, p. 158) tʃ <sup>h</sup> iu <sup>11</sup> -lim <sup>55</sup> / tʃ <sup>h</sup> iu <sup>11</sup> -nã <sup>55</sup> (Duffus, 1883, p. 114) kuaŋ ia (Dean, 1841, p. 18) 曠野 ts <sup>h</sup> iu <sup>11</sup> lim <sup>55</sup> (Fielde, 1883 p. 354) 樹林
			ST: ts <sup>h</sup> ru <sup>11</sup> lim <sup>55</sup> (Yeo, 2011, pg. 12)
			TT: lim <sup>55</sup> 林 (Atchariyasucha, 1982, pg. 121)
həu <sup>11</sup> ka <sup>11</sup> dəu <sup>11</sup>	give gift	kadou (gift) កាដូ	HT: loi <sup>53</sup> -mueŋ <sup>5</sup> / saŋ <sup>313</sup> -loi <sup>53</sup> / sie <sup>33</sup> -saŋ <sup>313</sup> (Duffus, 1883, p. 122) saŋ <sup>313</sup> -loi <sup>53</sup> -mueŋ <sup>5</sup> (Fielde, 1883, p. 465) 送禮物
			JT: li <sup>53-35</sup> mueŋ <sup>5</sup> (Xu, 2007, p. 180)

pi <sup>33</sup>	from (time or place)	pi: ពី	HT: taŋ <sup>11</sup> (Fielde, 1878, p. 172) 從 <sup>98</sup> t <sup>h</sup> aŋ <sup>313</sup> (Fielde, 1883, p. 576) 𠵼 ts <sup>h</sup> oŋ <sup>55</sup> / iu <sup>55</sup> (Ashmore, 1884, p. 55; Duffus, 1883, p. 117) ts <sup>h</sup> oŋ <sup>55</sup> (Goddard, 1888, p. 24; Lim, 1886, p. 5) 從
			JT: ts <sup>h</sup> oŋ <sup>55</sup> (Xu, 2007, p. 31)
			ST: taŋ <sup>11</sup> * <sup>99</sup> (Low, 2014, p. 54) t <sup>h</sup> aŋ <sup>213</sup> (Goh, 2017, p. 173) 𠵼
hai <sup>33</sup> , hei <sup>33</sup>	and (coordinator)	haəy ហើយ	
piŋ <sup>33</sup> pəŋ <sup>33</sup>	balloon	pe:ŋpaəŋ ប៉េងប៉ាំង	HT: p <sup>h</sup> u <sup>55</sup> -hun <sup>55</sup> -kiu <sup>55</sup> / t <sup>h</sup> iən <sup>33</sup> -tsun <sup>55</sup> / pue <sup>33</sup> -t <sup>h</sup> ia <sup>33</sup> / k <sup>h</sup> i <sup>313</sup> -kiu <sup>55</sup> (Duffus, 1883, p. 17)
			ST: be <sup>33</sup> loŋ <sup>55</sup> (Goh, 2017, p. 206) 码隆

Table 8.1 Khmer Loanwords in Cambodian Teochew

### 8.3 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), as shown in Chapter 4 Cambodian Teochew has only /l/. Therefore, Khmer /r/ is pronounced as /l/, as seen

<sup>98</sup> Fielde (1878) notes that this character is not of direct equivalence but is “of similar signification” (p. 8).

<sup>99</sup> \*Low (2014) did not include citation tones in her pitch superscripts. Since the one appearance of *taŋ*<sup>11</sup> ‘from’ in her paper was in a sandhi environment, according to her tone sandhi rules, the citation form could be either *taŋ*<sup>55</sup>, *taŋ*<sup>35</sup>, or *taŋ*<sup>11</sup>.

in the words for ‘jungle’, ‘prahok’, and ‘ice cream’. Additionally, while Khmer allows consonant clusters at syllable onset as seen in *prei* ‘jungle’, *prɔhɔk* ‘prahok’, *lhɔŋ* ‘papaya’, and *krɔla:n* ‘kralan’, these are not found in Teochew. To resolve this, the loanword could delete one of the two consonants or insert an epenthetic vowel (Miao, 2005, p. 105). In the first three instances, we see epenthetic vowel insertion between the two consonants to get *pɔ<sup>33</sup>lɔɪ<sup>52</sup>*, *pɔ<sup>33</sup>la<sup>33</sup>hɔk<sup>5</sup>*, and *lɔ<sup>11</sup>hɔŋ<sup>55</sup>*. For ‘kralan’, the /r/ is deleted to get *ka<sup>33</sup>laŋ<sup>33</sup>*. Cambodian Teochew also restricts consonants in syllable coda and /n/ is not found (see §4.1). We can see that the final sound of Khmer *krɔla:n* ‘kralan’ is pronounced /ŋ/ in the CT word *ka<sup>33</sup>laŋ<sup>33</sup>*. 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.<sup>100</sup> There is evidence that the Khmer loanwords in Cambodian Teochew undergo tone sandhi as the phrases in (1) show. Similarly, they provide an environment for tone sandhi such that the preceding words surface in their sandhi tones, as shown in (2).

- 1) a. ផ្លែឈូង            水  
           lɔ<sup>11</sup>hɔŋ<sup>55-11</sup>        tɕui<sup>52</sup>  
           papaya                water  
           ‘papaya juice’ [G1F2-PS]
- b. អាម៉ុក                鱼  
           a<sup>11</sup>mɔk<sup>5-2</sup>            hu<sup>55</sup>  
           amok                    fish  
           ‘amok fish’ [G3F3-PS]

---

<sup>100</sup> 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 in §4.3.

- 2) a. 炒                      𠵼𠵼𠵼  
 tɕ<sup>h</sup>a<sup>52-24</sup>              lək<sup>2</sup>laŋ<sup>5</sup>  
 stir-fry                  lok lak  
 ‘stir-fried beef cubes’ [G1F2-PS, G1F4-PS, G3F2-PS]
- b. 伊      □                  ប៉ឹងប៉ឹង  
 i<sup>33</sup>      nim<sup>24-11</sup>          piŋ<sup>33</sup>pəŋ<sup>33</sup>  
 3SG    hold                  balloon  
 ‘She holds the balloon.’ [G1F1-12]

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. As outlined in §6.2.1, in CT Enumeration Modification Constructions, ObjRs are typically counted and referred to with a classifier. The canonical word order is EC + Classifier + ObjR. The classifier used is dependent on the type of ObjR. My consultants provided evidence that Khmer loanwords in Cambodian Teochew in these constructions can be used with a specific classifier like 粒 *liap*<sup>5</sup>, used for round-ish things, as in (3), or with the linker word 个 *kaŋ*<sup>55</sup>, which can also be used with most things, as shown in (4).

- 3) a. 三      粒                  𠵼𠵼𠵼  
 sā<sup>33</sup>    liap<sup>5-2</sup>              lo<sup>11</sup>həŋ<sup>55</sup>  
 three   CLF.ROUND    papaya  
 ‘three papayas’ [G1F4-PS]
- b. 三      粒                  𠵼𠵼𠵼  
 sā<sup>33</sup>    liap<sup>5-2</sup>              maŋ<sup>11</sup>k<sup>h</sup>u<sup>2</sup>  
 three   CLF.ROUND    mangosteen  
 ‘three mangosteens’ [G1F4-PS]

- c. 两 粒 សារម៉ារី  
 ɲɔ<sup>24-11</sup> liap<sup>5-2</sup> sau<sup>33</sup> mau<sup>33</sup>  
 two CLF.ROUND rambutan  
 ‘two rambutans’ [G1F3-PS]
- 4) a. 三 个 ផ្លែឈូង  
 sã<sup>33</sup> kai<sup>55-11</sup> lo<sup>11</sup> hɔŋ<sup>55</sup>  
 three LW papaya  
 ‘three papayas’ [G3F1-PS, G3F2-PS]
- b. 三 个 មង្គុត  
 sã<sup>33</sup> kai<sup>55-11</sup> maŋ<sup>11</sup> k<sup>h</sup>u<sup>2</sup>  
 three LW mangosteen  
 ‘three mangosteens’ [G3F3-PS]
- c. 两 个 សារម៉ារី  
 ɲɔ<sup>24-11</sup> kai<sup>55-11</sup> sau<sup>33</sup> mau<sup>33</sup>  
 two LW rambutan  
 ‘two rambutans’ [G3F1-PS, G3F2-PS]
- d. 四 个 ប៉ឹងប៉ាំង  
 si<sup>11-53</sup> kai<sup>55-11</sup> piŋ<sup>33</sup> pɔŋ<sup>33</sup>  
 four LW balloon  
 ‘four balloons’ [G3F2-PS]

Khmer also has limited morphology, though one prefix used is *phlaε* meaning ‘fruit’. The word *phlaεbɔ̃* ‘avocado’ can be translated to ‘fruit-butter’. In isolation, *bɔ̃* is a French borrowing in Khmer meaning ‘butter’. The word *phlaε* can be used with fruits like ‘papaya’ (*phlaεlhɔŋ*), ‘mangosteen’ (*phlaεmɔŋkɨt*), and ‘rambutan’ (*phlaεsa:vmarv*). For these words, *phlaε* is not obligatory as the words without said prefix do not have another meaning, unlike with ‘avocado’ (*phlaεbɔ̃*) and ‘butter’ (*bɔ̃*). Importantly, when Cambodian Teochew borrows the Khmer word, *phlaε* is never used. The loanword for avocado becomes solely *bur*<sup>33</sup>. With

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

#### 8.4 Speaker Variation

It should be noted that not all of the eleven speakers interviewed for this study necessarily use all the words in Table 8.1. Table 8.2 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 §8.5. At least three speakers needed to use the loanword for it to be included in Table 8.1. ‘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*.

English Gloss	Word Used	Speaker
jungle, forest	7 pɔ̃ <sup>33</sup> lɔi <sup>52</sup>	G1F1, G1F2, G1M1, G2F1, G2M1, G2M2, G3F1
	1 pɔ̃ <sup>33</sup> lɔi <sup>52</sup> + tɕ <sup>h</sup> iu <sup>11</sup> lim <sup>55</sup>	G1F3
	1 tɕ <sup>h</sup> iu <sup>11</sup> lim <sup>55</sup>	G1F4
	2 no word	G3F2, G3F3
mangosteen	10 maŋ <sup>11</sup> k <sup>h</sup> u <sup>2</sup>	G1F1, G1F2, G1F3, G1F4, G2F1, G2M1, G2M2, G3F1, G3F2, G3F3
	1 unknown	G1M1

rambutan	10 sau <sup>33</sup> mau <sup>33</sup>	G1F1, G1F2, G1F3, G1F4, G2F1, G2M1, G2M2, G3F1, G3F2, G3F3
	1 unknown	G1M1
avocado	4 bu <sup>33</sup>	G1F1, G1F3, G1F4, G3F1
	1 gu <sup>11</sup> ni <sup>11</sup> kue <sup>52</sup>	G3F3
	1 gu <sup>11</sup> iu <sup>55-11</sup> kue <sup>52</sup>	G2M2
	1 a:vək <sup>h</sup> a:dou	G3F1
	3 no word	G1F2, G2F1, G2M1
	1 unknown	G1M1
papaya	10 lo <sup>11</sup> hoŋ <sup>55</sup>	G1F1, G1F2, G1F3, G1F4, G2F1, G2M1, G2M2, G3F1, G3F2, G3F3
	1 unknown	G1M
papaya salad	11 bək <sup>2</sup> lo <sup>11</sup> hoŋ <sup>55</sup>	all
prahok	11 pə <sup>33</sup> la <sup>33</sup> hək <sup>5</sup>	all
Cambodian crepe	11 bəŋ <sup>33</sup> tə <sup>h</sup> au <sup>52</sup>	all
ice cream	7 ka <sup>11</sup> lem <sup>33</sup>	G1F1, G1F2, G1F4 G2F1, G2M1, G3F1, G3F2
	1 siə <sup>2-5</sup> kə <sup>33</sup>	G3F3
	1 səŋ <sup>33</sup> kə <sup>33</sup>	G1F3
	2 unknown	G1M1, G2M2
kralan	3 ka <sup>33</sup> laŋ <sup>33</sup>	G1F4, G3F2, G3F3
	2 tek <sup>2-5</sup> kəŋ <sup>33</sup> kue <sup>52</sup>	G1F3, G2M2
	2 no word	G2F1, G3F1
	4 unknown	G1F1, G1F2, G1M1, G2M1
lok lak	8 lək <sup>2</sup> la <sup>25</sup>	G1F1, G1F3, G1M1, G2F1, G2M1, G2M2, G3F1
	3 tə <sup>h</sup> a <sup>52-24</sup> lək <sup>2</sup> la <sup>25</sup>	G1F2, G1F4, G3F2
	1 unknown	G3F3
amok	9 a <sup>11</sup> mək <sup>5</sup>	G1F1, G1F2, G1F3, G1F4, G2F1, G2M1, G2M2, G3F1, G3F3
	2 unknown	G1M1, G3F2
to give a gift	3 həu <sup>11</sup> ka <sup>11</sup> dəu <sup>11</sup>	G3F1, G3F2, G3F3
	2 saŋ <sup>53</sup> loi <sup>52</sup>	G2F1, G2M2

to gift something	2 saŋ <sup>53</sup> mueŋ <sup>22</sup> kiã <sup>24</sup>	G1F1, G1F3
to gift a book	1 saŋ <sup>55</sup> pəu <sup>24</sup>	G1F4
to give	3 həu <sup>11</sup>	G1F2, G1M1, G2M1
from	3 pi <sup>33</sup>	G3F1, G3F2, G3F3
	8 unknown	G1F1, G1F2, G1F3, G1F4, G1M1, G2F1, G2M1, G2M2
and	2 hai <sup>33</sup>	G1F1, G1F2
	1 hei <sup>33</sup>	G3F2
	5 tɕiu <sup>11</sup>	G1F4, G1M1, G2F1, G3F1, G3F3
	3 unknown	G1F3, G2M1, G2M2
balloon	6 piŋ <sup>33</sup> pəŋ <sup>33</sup>	G1F1, G2F1, G2M2, G3F1, G3F2, G3F3
	5 unknown	G1F2, G1F3, G1F4, G1M1, G2M1

Table 8.2 Breakdown of the count of speakers using each word

## 8.5 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, p. 46). The data in Tables 8.1 and 8.2 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 (G1F4, G3F2, G3F3) said the Khmer loanword *ka<sup>33</sup>laŋ<sup>33</sup>*. Two speakers (G2F1, G3F1) did not know a Teochew word for it and two (G1F3, G2M2), said 竹管糍 *tek<sup>2-5</sup>kəŋ<sup>33</sup>kue<sup>52</sup>* 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 竹 *tek<sup>2</sup>* ‘bamboo’ and 糍 *kue<sup>52</sup>* ‘cake’, rather than adopting the Khmer term.<sup>101</sup>

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 color with hair-like protuberances. Singapore Teochew’s *aŋ<sup>55</sup>mo<sup>55</sup>taŋ<sup>33</sup>* (Goh, 2017, p. 205) (and Mandarin’s 红毛丹 *hoŋ<sup>35</sup>mao<sup>35</sup>tan<sup>55</sup>*) was created with phono-semantic matching; *aŋ<sup>55</sup>mo<sup>55</sup>* ‘red hair’ describes the appearance of the fruit and the *taŋ<sup>33</sup>* 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 *mangis*. Singapore Teochew has again used phono-semantic matching with the Malay word in *maŋ<sup>55</sup>hek<sup>l</sup>* 芒黑 (Goh, 2017, p. 209), where *maŋ<sup>55</sup>* matches the source word’s pronunciation and *hek<sup>l</sup>* ‘black, dark’ describes the fruit’s outer color. 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<sup>33</sup>mau<sup>33</sup>* and *maŋ<sup>l</sup>k<sup>h</sup>u<sup>2</sup>* (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

---

<sup>101</sup> Thus *tek<sup>2-5</sup>kəŋ<sup>33</sup>kue<sup>52</sup>* is also an example of an expansive vocabulary item.

historical sources. It is likely a new food item introduced after Teochew speakers had arrived in Cambodia. The Khmer loanword *bur*<sup>33</sup>, attested in four speakers (G1F1, G1F3, G1F4, G3F1), was likely adopted out of efficiency. Alternate words include 牛奶果 *gu*<sup>55-11</sup>*ni*<sup>11</sup>*kue*<sup>52</sup> (‘milk fruit’) used by G3F3, and 牛油果 *gu*<sup>55-11</sup>*iur*<sup>55-11</sup>*kue*<sup>52</sup> (‘butter fruit’) used by G2M2. In these cases, existing Teochew words were put together to create the new word ‘avocado’.<sup>102</sup> The Khmer word ‘ice cream’ *ka:reem* is borrowed from French *crème*. Since the CT form *ka*<sup>11</sup>*lem*<sup>33</sup> 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 雪糕 *siɔ*<sup>2-5</sup>*kɔ*<sup>33</sup> which translates literally to ‘snow cake’. In contrast, Singapore Teochew uses the loanword 爱士吉林 *ai*<sup>33</sup>*sə*<sup>33</sup>*kik*<sup>1</sup>*lim*<sup>55</sup> (Goh, 2017, p. 205) or *ai*<sup>33</sup>*si*<sup>11</sup>*k*<sup>h</sup>*ɔ*<sup>11</sup>*lim*<sup>33</sup> (Chen, 2014, p. 856) and Johor Bahru Teochew uses *ai*<sup>33</sup>*si*<sup>33</sup>*k*<sup>h</sup>*i*<sup>2</sup>*lim*<sup>33</sup> (X. Chen, 2014, p. 840) from English ‘ice cream’.

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, CT speakers seemingly chose to adopt the Khmer word for ‘jungle, forest’ in *pɔ*<sup>33</sup>*lor*<sup>52</sup>, used by eight speakers in this study. Per Haspelmath (2009, p. 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 discern a difference between ‘forest’ and ‘jungle’. Evidence points to some usage of historic Teochew terms, in addition to the

---

<sup>102</sup> 牛油果 *gu*<sup>55-11</sup>*iur*<sup>55-11</sup>*kue*<sup>52</sup> ‘butter fruit’ could also be a loan translation from Khmer *phlaεbɔ* ‘fruit-butter’.

adoption of a Khmer loanword. As Table 8.2 shows, G1F3 used both *pɔ̃<sup>33</sup>lɔ̃<sup>52</sup>* and the historic *tɕ<sup>h</sup>iu<sup>11</sup>lim<sup>55</sup>*, while G1F4 used solely the latter term. G3F2 and G3F3 never produced a Teochew word for ‘jungle’ or ‘forest’, which hints at its rarer usage in Cambodia.<sup>103</sup> Khmer ព្រៃ *prei* ‘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<sup>55</sup>* 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.

Cambodian Teochew *pi<sup>33</sup>* ‘from’, which comes from Khmer ពី *pi:* is an example of the borrowing of a function word.<sup>104</sup> Generally, we can see *taŋ<sup>11</sup>* ‘from’ used in time expressions in Singapore Teochew (5a) and historical Teochew (5b), and in locative expressions in historical Teochew in (6).

---

<sup>103</sup> 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, G3F1 used a loanword *bəŋ<sup>33</sup>* ‘lake’ from Khmer បឹង *bəŋ* (versus historic Teochew 湖 *ɔu<sup>55</sup>* used by three speakers, G1F3, G3F2, G3F3). Meanwhile, G3F2 and G3F3 used a loanword *ap<sup>5</sup>* for ‘fog’ from Khmer អង្គ *lap* (versus historic Teochew terms 蒙 *məŋ<sup>55</sup>*, 霧 *məu<sup>55</sup>*, and 雾 *bu<sup>11</sup>* used by G1F4, G1F3, and G3F1 respectively).

<sup>104</sup> 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 *ta<sup>33</sup>pi<sup>5</sup>* ‘but’ in Singapore Teochew (Yeo, 2011), all borrowed from Malay.

- 5) a. a<sup>33</sup>-Zach    taŋ<sup>11</sup>    sɔi<sup>213</sup>    tsu<sup>11</sup>    tsɨŋ<sup>33</sup>    tsĩa<sup>11</sup>gau<sup>11</sup>    pua<sup>ʔ2</sup>    lai<sup>55</sup>    pua<sup>ʔ2</sup>    khi<sup>11</sup>  
 NVOC-PN    from    young    SEQ    very    good.at    fall    come    fall    go  
 ‘From a young age, Zach was very accident-prone.’ (Low, 2014, p. 54)<sup>105</sup>
- b. taŋ<sup>11</sup>    ts<sup>h</sup>u<sup>ʔ2</sup> si<sup>313</sup>    kau<sup>313</sup>    tã<sup>33</sup>  
 from    birth    arrive    now  
 ‘from his birth till now’ (Fielde, 1878, p. 172)
- 6) taŋ<sup>11</sup>    tsi<sup>53</sup>    siã<sup>55</sup>    kau<sup>313</sup>    nin<sup>53</sup>    tse<sup>11</sup>    lai<sup>25</sup>    u<sup>25</sup>    dzie<sup>ʔ55</sup>    tsoi<sup>25</sup>    li<sup>53</sup>    lo<sup>25</sup>  
 from    this    city    arrive    2PL    village    there is    how    much    distance  
 ‘How far is it from this city to your village?’ (Fielde, 1878, p. 240)

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

- 7)    មក    ពី    ស្រុក    អង់គ្លេស  
 mɔk    pi:    srok    ʔaŋkleeh  
 come    from    country    England  
 ‘(She) comes from England.’ (Sak-Humphry, 2016, p. 5)
- 8)    ចាប់    ពី    ម៉ោង    ប្រាំ    កន្លះ    រហូតដល់    ម៉ោង    ដប់    យប់  
 cap    pi:    moaŋ    pram    kanlah    rɔhoot    dal    moaŋ    dap    yup  
 starting    from    hour    five    half    until    hour    ten    night  
 ‘starting from 5:30PM until 10:00PM’ (Sak-Humphry, 2016, p. 50)

There is evidence that loanword *pi*<sup>33</sup> ‘from’ in Cambodian Teochew is similarly used with both locative (9) and time expressions (10).

---

<sup>105</sup> In Low (2014), NVOC is a non-vocative particle, PN means proper name, and SEQ is a sequentiality marker.

9) 伊 爬 出 𠵼 房  
*i*<sup>33</sup> *pe*<sup>55-11</sup> *tɕ*<sup>h</sup>*uk*<sup>2-5</sup> *pi*<sup>33</sup> *paŋ*<sup>55</sup>  
 3SG crawl exit from room  
 ‘He crawled out from his room.’ [G3F1-2]

10) 𠵼 一 点 到 五 点  
*pi*<sup>33</sup> *ik*<sup>2-5</sup> *tiam*<sup>52</sup> *kau*<sup>11-53</sup> *ŋɔu*<sup>24-11</sup> *tiam*<sup>52</sup>  
 from one o'clock arrive five o'clock  
 ‘from 1 o'clock to 5 o'clock’ [G3F2-E]

Evidence shows that Cambodian Teochew has replaceive borrowings in ‘papaya’ and ‘balloon’. The word 木瓜 *bak<sup>5</sup>kue<sup>33</sup>* ‘papaya’ was found in Fielde (1883, p. 277) and in Indonesian and Singapore Teochew varieties. No evidence was found for the existence of *bak<sup>5</sup>kue<sup>33</sup>* in Cambodian Teochew. Conversely, *lɔ<sup>11</sup>hɔŋ<sup>55</sup>* ‘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, *bɔk<sup>2</sup>lɔ<sup>11</sup>hɔŋ<sup>55</sup>*. Since the Cambodian Teochew were already using *lɔ<sup>11</sup>hɔŋ<sup>55</sup>* in this dish’s name, it would be more efficient to refer to the fruit ‘papaya’ using the same word. Duffus (1883, p. 17) presented several words for ‘balloon’ including *p<sup>h</sup>u<sup>55</sup>-hun<sup>55</sup>-kiu<sup>55</sup>*, *t<sup>h</sup>iɛn<sup>33</sup>-tsun<sup>55</sup>*, and *k<sup>h</sup>i<sup>313</sup>-kiu<sup>55</sup>*. Each of these words has the morpheme 球 *kiu<sup>55</sup>* ‘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 (G1F1, G2F1, G2M2, G3F1, G3F2, G3F3) and no evidence was found for the use of historic Teochew words. Singapore Teochew has also adopted a loanword 码隆 *be<sup>33</sup>loŋ<sup>55</sup>* from English (Goh, 2017, p. 206).

Finally, Cambodian Teochew has one loan blend in ‘to give a gift’. This is a hybrid Teochew-Khmer phrase in which the predicate ‘give’ is pronounced as Teochew 𠵿 *hɔu<sup>11</sup>*, while the ObjR ‘gift’ *ka<sup>11</sup>dɔu<sup>11</sup>* comes from Khmer កាដូ *kadov* ‘gift’; *hɔu<sup>11</sup>ka<sup>11</sup>dɔu<sup>11</sup>* is specifically used to mean to give someone a gift or present. No evidence was found of *ka<sup>11</sup>dɔu<sup>11</sup>* being used in other contexts and it always appeared with *hɔu<sup>11</sup>*. 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 *hɔu<sup>11</sup>* when shown the prompt; *hɔu<sup>11</sup>* ‘to give’ is used in Cambodian Teochew in phrases such as ‘to give money’ or ‘to give plants’. Five speakers used a phrase with 送 *saj<sup>53</sup>*, which means ‘to gift’.<sup>106</sup> Two speakers said 送礼 *saj<sup>53</sup>lɔi<sup>52</sup>* ‘to give a gift’, two 送物件 *saj<sup>53</sup>mue<sup>2</sup>kiã<sup>24</sup>* ‘to give something’ and one 送簿 *saj<sup>53</sup>pɔu<sup>24</sup>* ‘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.

Fourteen of the sixteen words found in this chapter and presented in Table 8.1 are content words versus function words and these fourteen are all ObjRs rather than predicates. 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.<sup>107</sup> However, the sixteen words do not constitute an exhaustive list of Cambodian Teochew’s Khmer loanwords. Another 50 or so Khmer words have been identified as loanwords in Cambodian Teochew by X. Chen (2014, pp. 915-18). More examples would likely surface with additional fieldwork and data analysis. Because the methodology relied on

---

<sup>106</sup> Citation tone for this word in Cambodian Teochew is unconfirmed.

<sup>107</sup> As opposed to core vocabulary (cf. Myers-Scotton, 1993; Haspelmath, 2009).

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.<sup>108</sup>

---

<sup>108</sup> 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.

## Chapter 9 – Conclusion

This dissertation concludes with a summary of the key findings presented in the prior eight chapters (§9.1), an overview of the limitations of the current study and areas for improvement and future research (§9.2), and finally some implications of the current work (§9.3).

### 9.1 Summary

In Chapter 1, we provided some background information on the Teochew language family, as well as outlined the framework and methodology for the current study. Chapter 2 highlighted the sociolinguistic situation of the Teochew in Cambodia, as well as in other regions of Southeast Asia. Despite the historic Sinitic influence on local languages like Khmer and Thai, as well as on cultural practices in the region, there is a language shift happening in the region amongst minority speakers as they move away from languages like Teochew in favor of Mandarin, English, and the respective country's majority language, such as Khmer. Chapter 3 showcased some of the prior research that has been done on Chaoshan varieties. Much of the work has focused on varieties in China, though some researchers have studied other Southeast Asian varieties. The Cambodian Teochew variety is severely understudied and this work represents the first detailed grammar of this language. Chapter 4 covered the Cambodian Teochew phonology. While the sound system displays similarities with other Teochew varieties including the use of contrastive vowel nasalization, tone, and tone sandhi, the extent of these phenomena and patterns of use are somewhat distinctive. Cambodian Teochew also makes use of alveolo-palatal affricates as compared to other varieties' alveolar affricates, though younger speakers are showing deaffrication in the voiced alveolo-palatal

affricate, /dz/. Younger speakers are also showing reduced tonal distinctions in sandhi environments.

Chapters 5, 6, and 7 moved into the Radical Construction Grammar portion of the grammar, focusing on the three propositional act functions of reference, modification, and predication, respectively. Each chapter covered how the semantic classes of objects, properties, and actions participated in each function.

In Chapter 5, we showed a potential simplification in the system of CT Personal Contextual ObjR Constructions in the lack of distinctive CT Inclusive and Exclusive First Person Contextual ObjR Constructions, with some speakers using 俺 *naŋ*<sup>52</sup> ‘1PL’ for both contexts, and the use of 伊 *i*<sup>33</sup> for both CT Third Person Singular and Plural Contextual ObjR Constructions. This chapter also explored how the language treats referents in regards to their information status. We showed the frequent use of CT Zero Anaphora and CT Anaphoric Contextual ObjR Constructions for active referents, as well as the use of 伊 *i*<sup>33</sup> ‘IK’, “bare” classifiers, and 个 *kaŋ*<sup>55</sup> ‘IK’ as CT Identity Known Constructions for semi-active referents. We also illustrated the use of CT Interrogative Contextual ObjRs as CT “indefinite” Contextual ObjRs (Haspelmath, 1997). Finally, CT Action Reference Constructions make frequent use of the CT Action Predicate 𠵼 *tã*<sup>21</sup> ‘say’ that has grammaticalized into a complementizer, seemingly following the pathway from Chappell (2008).

In Chapter 6, due to the relatively few forms of unique classifiers used in the data set and the prevalence of the use of the generic classifier 个 *kaŋ*<sup>55</sup>, we have argued for a simplification of the classifier system in Cambodian Teochew. There is also some evidence that the classifier is not strictly required for CT Enumeration Modification Constructions or CT Deictic Attributive Constructions. We also showed some unexpected word orders in CT

Enumeration Modification Constructions, as well as post-ObjR modification with quantifier 𑜁𑜃𑜂𑜫 *tɕɔi<sup>11</sup>* ‘many’, and some CT Action Modification Constructions using the post-objR strategy. These word orders differ from other Teochew varieties and show an increasing tendency for post-ObjR modification, which is also found in Khmer. Thus we hypothesize that contact with Khmer is causing these word order changes in Cambodian Teochew.

Chapter 7 highlighted the most common CT Constructions fulfilling the predication subfunction, of which there are many. We argued for further simplification in the grammar in terms of CT Declarative Negation Constructions, which seem to be showing a preference for the use of 𑜁𑜃𑜂𑜫 *bɔ<sup>55</sup>* ‘NEG’. We showed numerous examples of CT Complex Predicate Constructions, which make frequent use of the serial predicate strategy, as is common in Sinitic languages. CT Property Predication Constructions frequently arise with referent + predicated property and no additional modification of the property like an intensifier. This is arguably counter to what arises in Mandarin (Sun, 2022, pp. 208-9), Cantonese (Francis & Matthews, 2005, p. 277), or Jieyang Teochew (Xu, 2007, p. 52), and may show influence from Khmer where conversely and unmodified predicated property is completely acceptable. In using the copula strategy for CT Object Predication Constructions and not CT Action and Property Predication Constructions, Cambodian Teochew follows the universals from Croft (1991, p. 130) and Stassen (1997, p. 127). Finally, CT Locative Predication Constructions seem to show an influence from Khmer in that the path word frequently precedes the ground, as it does in Khmer.

Chapter 8 provided a preliminary list of Khmer words that have been integrated into the Cambodian Teochew language. The data provided evidence for expansive, additive, replacive, and loan blend loanwords in Cambodian Teochew. The majority of the words were content

words versus function words. This initial list of sixteen borrowings in Table 8.1 is considerably shorter than comparative lists of loanwords in other Teochew varieties (Chen, 2014; Goh, 2017; Lin, 2006). More research needs to be done to see if this is due to the limited data in the current study or due to an underlying resistance to lexical change.

## 9.2 Limitations and future work

The present study had several limitations and thus many areas for improvement and expansion in future research.

The current study was limited in the number of speakers that were interviewed. We have shown that language variations exist even amongst the nine primary consultants and six secondary consultants. More variation would likely be found in a larger sample size. We have hinted at some patterns that may be related to a speaker's generation, such as the differing yin shang, high falling tone, in sandhi environments for G3 speakers. However, data from more G3 speaker consultants is needed to help us determine the extent of this pattern. As it stands, given the relatively few number of speakers consulted in this project, we are not able to do any type of variationist work, such as comparing patterns of language use among differing populations of people such as by gender, education level, age/generation, etc. That type of sociolinguistic analysis would require many more speakers.

Improvements could have been done in the selection of speakers that we did interview. Seven of the speakers (five primary and two secondary) in this project can be traced to the same family, though they represent four different households. Thus there is the potential that the data shows the Teochew spoken by only one family. Attempts made to diversify participants by recruiting from online spaces did succeed in bringing in an additional four primary consultants, all from separate families. Yet the online recruiting process itself likely

influenced the type of speakers integrated in this project. Online spaces included the 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 possible that their Teochew had degraded due to their use of other languages and thus their usage of Cambodian Teochew might differ from others in the community.

Outside of the primary consultants, the secondary texts provided data from four other speakers which also helped in creating a corpus from a diverse bunch.

Even with the data from fifteen total speakers from nine families, it is impossible to generalize about Cambodian Teochew as a whole and this dissertation attempted to only present findings based on the given data set. Additionally, though we have called this variety Cambodian Teochew, all speakers interviewed were from Phnom Penh. There is no data on the language spoken by Cambodian Teochews living outside the capital city. There is likely variation amongst speakers from different regions in Cambodia, similar to Peng's (2012) findings on the Jambi and Pontianak varieties of Teochew in Indonesia. Consultants outside of Phnom Penh would be needed in order to investigate the existence of language varieties within Cambodia itself.

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, since speakers were aware of my research intent, they may

have presented to me a certain way of speaking that differed from their normal everyday speech. 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. In the picture stimuli sessions, picture prompts were used 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 and this may have affected the speakers' responses. Similarly, in the picture book narration task, sometimes the interpreter would start narrating the story or describing the picture themselves in Khmer or Teochew. A few times there was also some confusion in consultants as to purpose of the task due to miscommunication on my part leading to unclear instructions from the interpreters. As a result of this, some speakers were describing items on a page in detail, rather than putting the pictures together in the context of a story.

The selection of stories to narrate also could have been improved. The *Zoom and While You Are Sleeping* books lacked a cohesive story and therefore lent themselves to more descriptive speech. Other stories narrated by primary consultants such as *Looking for Felix*, *Cat Story*, *Frog*, *Where Are You?*, *The Red Book*, and *Fann* were primarily focused on the story/experience of one actor, rather than the interactions and dialogues between multiple actors. Therefore, the types of constructions that were found in these texts were limited. The text that lent itself to the widest variety of constructions was the family problem picture task from San Roque et al. (2012). Though the family problem task provided an abundance of great data, it was not actually as successful as I had hoped. In addition to consultants literally describing the pictures and telling the story, the task is designed to elicit speech between consultants as they discuss what they think is happening in the pictures, and arrange them in an

appropriate order for the narrative. However, in this case, the speech between consultants was in Khmer, and Teochew was reserved for the actual picture-describing. This says something about the consultants' (G3F2 and G3F3) preferences for language use. While they both have shared languages in English, Khmer, Mandarin, and Teochew, they chose to use Khmer to accomplish this task, even though I (unsuccessfully) encouraged them to converse in Teochew. It would be beneficial to repeat this task with other groups/pairs of Cambodian Teochew speakers to see their language of choice for purposes other than the picture description part of the task.

One big benefit of using the secondary data was that it provided different genres of texts, and thus increased the types of constructions found in the current data set. The garden tour text offered more of a presentational text, on a topic that the speaker was excited and passionate about. It also provided specialized vocabulary from vegetation terminology. The news interview had examples of more personal stories from speakers, which lead to the use of more CT Personal Contextual ObjR Constructions, for example. From speaker G2F1US, *The Flower*, *The Little Carp that Jumped over the Dragon Gate*, *The Little Snail*, and *The Smart Fox and the Tiger* were more interactive stories, showcasing dialogue between speakers, and thus included examples of CT Interrogative Constructions, and CT Imperative-Hortative Constructions which were otherwise lacking in the primary data. The use of these stories was not without its flaws, however. The methodology for getting these texts was that speaker G2F1US would listen to a page of the story in Mandarin and then retell that portion of the story in Teochew. This procedure lent itself to language interference from Mandarin. In the future, it would be good to strike a balance between the more descriptive picture stories in the primary

data and the stories that required narration in another language such as Mandarin, in the secondary data.

Though the secondary data provided more examples of constructions not found in the primary data set, using that data itself was flawed because of the background of the speakers. For the three speakers in the news interview, it's clear that they resided in Cambodia at the time of filming. However, because I was not able to interview them in the same way as my primary consultants, we have no idea of their family background and language use. Thus it cannot be confirmed when their families immigrated to Cambodia, when they started speaking Teochew, and in what domains they speak it. As for the other secondary speakers, G2F1US, G2M2US and G1F5US, as their identifiers indicate, they have lived in the United States for decades. Thus their language could be different from speakers who currently reside in Cambodia. Language change could have resulted in their Cambodian Teochew as a result of contact from English, or through reduced domains of use, and language attrition.

At the same time, the speakers living in Cambodia themselves have not escaped a reduction in language use. The primary speakers have reported primarily only using the language in the home domain. Due to the age of speakers G1F3 and G1F4, they noted that they do not speak the language often anymore because many of their friends who they used to speak it with are no longer around, and their younger family members do not speak much at all. Thus we are also likely seeing the effects of language attrition for the primary speakers as well.

Throughout this grammar we have noted areas that could be improved upon in future research. Due to a lack of data, descriptions of many different constructions were absent altogether or would benefit from additional data. Future research should focus on gathering data for the following construction types: CT Personal Contextual ObjR Constructions, CT

Modification Constructions that show part-whole relations, CT Interrogative Constructions including alternative questions, CT Response Constructions, CT Comparative Constructions, CT Nonbasic Voice Constructions, CT Imperative-Hortative Constructions, CT Action Modification Constructions, CT Property Modification Constructions including those that use admodifiers, CT Possession Constructions, CT Locative Modification Constructions, CT Declarative Negation Constructions with a focus on different types of predicates, CT Discourse Markers, and many more. Ultimately this dissertation cannot be considered a full reference grammar as the findings are preliminary and in no way comprehensive. Ideas for future procedures include encouraging the telling of personal narratives and dialogues between speakers.

Certain aspects of the phonology could also be explored in more depth such as patterns of tone sandhi. For example, it was shown in Table 5.3 that in CT Contextual ObjR Constructions formed with 只 *tɕi<sup>52</sup>* ‘this’ and 许 *hu<sup>52</sup>* ‘that’, said syllables seemingly do not undergo tone sandhi, despite being in a position where tone sandhi would be expected. Xu (2007) reported the Jieyang Deictic Contextual ObjR Constructions as *tsi<sup>53-35</sup>ka<sup>55</sup>* ‘PDCO’ and *hia<sup>53-35</sup>ka<sup>55</sup>* ‘DDCO’, whose tone sandhi on the first syllable is evident. Conversely, the CT Deictic Contextual ObjR Constructions, as reported by my consultants, are 只个 *tsi<sup>52</sup>ka<sup>55</sup>* ‘PDCO’ and 许个 *hu<sup>52</sup>ka<sup>55</sup>* ‘DDCO’. Other areas where tone sandhi was expected yet not always present was in CT Transitive Constructions (see §7.1) and CT Complex Predicate Constructions (see §7.1.4). While I have occasionally pointed out some patterns of tone sandhi in this grammar of Cambodian Teochew, it has never been my focus. Future studies could delve deeper in this topic.

We have noted when some variations in speaker pronunciation were found, like in 新 *siŋ*<sup>33</sup> vs. *seŋ*<sup>33</sup> ‘new’. The *seŋ*<sup>33</sup> pronunciation can be traced back to the Jieyang variety, according to Mogher. Thus this variation in Cambodian Teochew may exist as a result of the place of origin in China of particular speakers’ families. However, given the length of time these families had already lived in Cambodia, unfortunately many consultants were unaware of their family’s origin. If one is able to find a sufficient number of speakers who do have an idea of the Chaoshan variety their ancestors spoke, future work could look at how the differing pronunciations from these varieties have manifested in Cambodian Teochew. This has been done in regards to tone in Khoo (2018)’s study on the fishing village in Parit Jawa in Johor Malaysia, whose descendants from China hailed with three different Teochew varieties.

We briefly noted instances from speaker G1F1 who seems to be showing a sound change in process from /g/ → /w/ → /b/. Future research could study the extent of this sound change for G1F1, and see if the change is similarly in process with other speakers.

This paper did not extensively cover the CT lexicon. While Chapter 8 showed the usage of some Khmer loanwords, throughout this dissertation we have also seen borrowings/codeswitching with words from Mandarin and English. Future research should look at the use of these languages in Cambodian Teochew. It is especially important to study loanwords in languages of the Chinese diaspora, as the integration of local words in the Sinitic lexicon is practically a defining feature of all overseas varieties (Zhang, 2013). We covered lexical innovations by speakers in CT Object Modification Constructions in §6.3.2 and in CT Action Modification Constructions in §6.4. It would be interesting to talk with speakers about their choice in using those types of innovations/constructions instead of code-switching/borrowing a word from Khmer, English, or Mandarin. It would also be interesting to

look at how widespread/lexicalized any of the given words were (for the same speakers and for other speakers), or if they were merely an in-the-moment type of language production.

The conclusions in this paper would benefit from more conversations with speakers on their language use, as we just mentioned. Throughout this paper we have mostly reported on what was found in the data rather than on things that speakers said about their language use. Improvements could be made by asking speakers to elaborate on certain language behaviors and phenomena.

At many points throughout this dissertation, we made reference and comparisons to constructions in other Teochew varieties, relying on the existing literature on said varieties, such as Xu (2007) and her other works for Jieyang, Goh (2017, Low (2014) and Yeo (2011) for Singapore Teochew, Peng (2011, 2012) for the Indonesian varieties, X. Chen (2003) for the Johor Bahru variety of Malaysia, X. Chen (2019) for three Thai varieties, and X. Chen (2014) for seven varieties throughout Southeast Asia (from Cambodia, Indonesia, Laos, Malaysia, Singapore, Thailand, and Vietnam). This method is not without its flaws as the data we pulled from the other studies will ultimately absorb the limitations of said works. For example, the aforementioned studies typically used data from only a few speakers, sometimes relied on elicitation, and sometimes lacked tone or tone sandhi markings. Many are also all eight or more years old. Therefore, in the same way we tried not to say that the data we found was representative of a Cambodian Teochew variety as a whole, we cannot say that the data found in the aforementioned studies necessarily represents each variety as a whole. Future research could try to obtain similar types of naturalistic data from the different varieties and make comparisons based on that data.

### 9.3 Implications

Despite the limitations of this study, it still has important implications for disciplines such as Cambodian and Sinitic linguistics, as well as contact linguistics and construction grammar.

In addition to the other Chinese groups in Cambodia, including those who speak Cantonese, Hakka, and Hokkien, there are several other minority groups, and thus minority languages spoken across the country which have similarly been understudied. This dissertation adds to the literature on the minority languages of Cambodia.

Additionally, since Sinitic languages of the Chinese diaspora are numerous, this study represents a good point of comparison for the potential of such languages when they are under pressure from a majority language, which is often the case.

We have added to the field of contact linguistics by showing what is possible in contact environments. Khmer influence was hypothesized in all aspects of the Cambodian Teochew grammar including the phonology, lexicon, and syntax. The evidence points to an overall simplification of the CT grammar, which is similar to the findings from other contact environments (Dorian, 1981; Gruzdeva, 2004, 2015; Kaysina, 2015; Romaine, 2010; Schmidt, 1983).

This dissertation is one of the first grammars written in a radical construction framework (Croft, 2001, 2013, forthcoming). While not a comprehensive grammar, it can still be used as a sample to show readers how one would go about writing a grammar in this manner. It also provides some possible terms for those who are looking to avoid resorting to the use of traditional terminology.

Finally, this study has implications for the Cambodian Teochew speakers themselves and their families, many of whom expressed excitement and gratitude at the idea of a study

done on their language. This work and future studies can emphasize to these speakers the importance of their language and intergenerational transmission.

## References

- Aikhenvald, Alexandra Y., & Dixon, R. M. W. (Eds.). (2006). *Serial verb constructions: A cross-linguistic typology*. Oxford University Press.
- Ann, Sovatha. (2011). Paper money in Phnom Penh: Beyond the Sino-Khmer tradition. *Explorations: A Graduate Student Journal of Southeast Asian Studies*, 11(1), 93-104.
- Arcodia, Giorgio Francesco. (2007). Chinese: A language of compound words? In F. Montermini, G. Boyé, & N. Hathout (Eds.), *Selected proceedings of the 5th décebrettes: Morphology in Toulouse* (pp. 79-90). Cascadilla Proceedings Project.
- Ariel, Mira. (1990). *Accessing noun phrase antecedents*. Routledge.  
<https://doi.org/10.4324/9781315857473>
- Ashmore, William. (1884). *Primary Lessons in Swatow Grammar*. English Presbyterian Mission Press.
- Atchariyasucha, Wichet. (1982). *A phonological study of Swatow of Chinese as spoken in Bangkok with comparisons to Thai* [Master's thesis, Mahidol University]. Mahidol University Library and Knowledge Center. <http://mulinet11.li.mahidol.ac.th/e-thesis/scan/14050.pdf>
- Aw, Tash. (2019, November 26). Coming out of the shadows: what it means to be French and Chinese. *The Guardian*. <https://www.theguardian.com/world/2019/nov/26/what-you-hear-about-chinese-people-in-france-feeling-scared-its-true>
- Bao, Zhiming. (1999). Tonal contour and register harmony in Chaozhou. *Linguistic Inquiry*, 30(3), 485-493. <https://www.jstor.org/stable/4179074>
- Bauer, Robert S. (2018). Cantonese as written language in Hong Kong. *Global Chinese*, 4(1), 103-142. <https://doi.org/10.1515/glochi-2018-0006>

- Birnie-Smith, Jessica Rae. (2016). Ethnic identity and language choice across online forums. *International Journal of Multilingualism*, 13(2), 165-183.  
<https://doi.org/10.1080/14790718.2015.1078806>
- Bybee, Joan L. (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins Publishing. <https://doi.org/10.1075/tsl.9>
- Cai, Hong. (2014). *The aspectual system in Gong'an Hua, a dialect of Southwestern Mandarin spoken in Hubei Province, China* [Doctoral dissertation, University of Queensland]. UQ eSpace. <https://espace.library.uq.edu.au/view/UQ:337463>
- Cai, Junming (蔡俊明). (1991). 潮州方言词汇 (*Lexicon of the Chaoshan dialects*). The Chinese University of Hong Kong Press.
- Chan, Sambath. (2005). *The Chinese minority in Cambodia: Identity construction and contestation* [Master's thesis, Concordia University]. Spectrum Research Repository. <https://spectrum.library.concordia.ca/8439/>
- Chao, Yuenren. (1968/2011). *A grammar of spoken Chinese*. University of California Press.
- Chappell, Hilary. (2001). Language contact and areal diffusion in Sinitic languages. In A. Y. Aikhenvald & R. M. V. Dixon (Eds.), *Areal diffusion and genetic inheritance: problems in comparative linguistics* (pp. 328-357).
- Chappell, Hilary. (2008). Variation in the grammaticalization of complementizers from *verba dicendi* in Sinitic languages. *Linguistic Typology*, 12, 45-98.  
<https://doi.org/10.1515/LITY.2008.032>
- Chappell, Hilary. (2015). *Diversity in Sinitic languages*. Oxford University Press.  
<https://doi.org/10.1093/acprof:oso/9780198723790.001.0001>
- Chappell, Hilary, & Creissels, Denis. (2019). Topicality and the typology of predicative possession. *Linguistic Typology*, 23(3), 467-532. <https://doi.org/10.1515/lingty-2019-0016>

- Chen, Chuanjia (陈传佳). (1996). 潮汕方言的宾语前置 (Preposition of objects in Chaoshan dialects). *韩山师范学院学报*(1), 98-104.
- Chen, Ee San. (2003). Language convergence and bilingual acquisition: the case of conditional constructions. *Annual Review of Language Acquisition*, 3(1), 89-137.
- Chen, Jifan (陈基藩). (1987). 潮汕方言助词“哩罗”的意义特点 (Semantic properties of the particles ‘li’ and ‘lo’ in Chaoshan dialect). *韩山师专学报 (社会科学版)*(2), 96-98.
- Chen, Shihlun Allen. (2015). *Socializing Chineseness: Cambodia's ethnic Chinese communities as a method* (Publication Number 10002238) [Doctoral dissertation, University of Hawai'i at Manoa]. ProQuest Dissertations & Theses Global.
- Chen, Weirong. (2011). *The Southern Min dialect of Hui'an: Morphosyntax and grammaticalization* [Doctoral dissertation, The University of Hong Kong]. The HKU Scholars Hub. <http://hub.hku.hk/handle/10722/141944>
- Chen, Xiaojin (陈晓锦). (2003). 马来西亚的三个汉语方言 (*The three Chinese dialects in Malaysia*). China Social Sciences Press.
- Chen, Xiaojin (陈晓锦). (2009). 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, 79-84.
- Cheng, Chin-Chuan, & Wang, William S-Y. (1977). Tone change in Chao-zhou Chinese: A study in lexical diffusion. In W. S. Wang (Ed.), *The lexicon in phonological change* (pp. 86-100). De Gruyter Mouton. <https://doi.org/https://doi.org/10.1515/9783110802399.86>
- Cheng, Lisa L.-S., & Sybesma, Rint. (2005). Classifiers in four varieties of Chinese. In G. Cinque & R. S. Kayne (Eds.), *The Oxford handbook of comparative syntax* (pp. 259-292). <https://doi.org/10.1093/oxfordhb/9780195136517.013.0007>
- Chin, James K. (2017). Ethnicized networks and local embeddedness: The new Chinese migrant community in Cambodia. In M. Zhou (Ed.), *Contemporary Chinese diasporas* (pp. 187-206). Springer. [https://doi.org/10.1007/978-981-10-5595-9\\_9](https://doi.org/10.1007/978-981-10-5595-9_9)

- Chokkajitsumpun, Pranee. (1998). *Chinese literacy maintenance and shift in Bangkok: Individual and family cases* [Unpublished doctoral dissertation, University of Hawai'i at Mānoa].
- Clayton, Thomas. (2006). *Language choice in a nation under transition: English language spread in Cambodia*. Springer. <https://doi.org/10.1007/0-387-31194-7>
- Cole, Peter, Hermon, Gabriella, & Lee, Cher Leng. (2001). Grammatical and discourse conditions on long distance reflexives in two Chinese dialects. In P. Cole, G. Hermon, & C.-T. J. Huang (Eds.), *Long distance reflexives* (pp. 1-46). Brill. [https://doi.org/10.1163/9781849508742\\_002](https://doi.org/10.1163/9781849508742_002)
- Cole, Peter, & Lee, Cher Leng. (1997). Locality constraints on yes/no questions in Singapore Teochew. *Journal of East Asian Linguistics*, 6(2), 189-211. <https://doi.org/10.1023/A:1008237705178>
- Comrie, Bernard. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge University Press.
- Cooper, Doug. (2020). *Notes on Modern Chinese loanwords in Thai*. Retrieved June 1 from <http://sealang.net/thai/chinese/modern.htm>
- Croft, William. (1990). A conceptual framework for grammatical categories (or: a taxonomy of Propositional Acts). *Journal of Semantics*, 7(3), 245-279. <https://doi.org/10.1093/jos/7.3.245>
- Croft, William. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. University of Chicago Press.
- Croft, William. (2001). *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198299554.001.0001>

- Croft, William. (2003). *Typology and universals* (2nd ed.). Cambridge University Press.  
<https://doi.org/10.1017/CBO9780511840579>
- Croft, William. (2005). Logical and typological arguments for Radical Construction Grammar. In J.-O. Östman & M. Fried (Eds.), *Construction Grammars: Cognitive grounding and theoretical extensions* (pp. 273–314). John Benjamins Publishing Company.  
<https://doi.org/10.1075/cal.3.11cro>
- Croft, William. (2007). The origins of grammar in the verbalization of experience. *Cognitive Linguistics*, 18(3), 339-382. <https://doi.org/10.1515/COG.2007.021>
- Croft, William. (2013). Radical Construction Grammar. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 211-232). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195396683.013.0012>
- Croft, William. (2016). Comparative concepts and language-specific categories: Theory and practice. *Linguistic Typology*, 20(2), 377-393. <https://doi.org/10.1515/lingty-2016-0012>
- Croft, William. (2020a). *Ten lectures on construction grammar and typology*. Brill.  
<https://doi.org/10.1163/9789004363533>
- Croft, William. (2020b). Word classes in Radical Construction Grammar. In E. V. Lier (Ed.), *The Oxford Handbook of Word Classes* (pp. TBD). Oxford University Press.
- Croft, William. (forthcoming). *Morphosyntax: Constructions of the world's languages*. Cambridge University Press.
- Cun, Xi. . (2009). *A phonetic study on implosives in China* (Publication Number 3365904) [Doctoral thesis, The Hong Kong University of Science and Technology]. ProQuest Dissertations & Theses Global. <https://search.proquest.com/docview/305151964>
- Cun, Xi. (2010). The phonetic cause of sound change from voiceless stops to implosives. *Bulletin of Chinese Linguistics*, 4(1), 33-65.
- Dahl, Östen. (1985). *Tense and aspect systems*. Blackwell.

- Dahles, Heidi, & ter Horst, John. (2012). Institutionalising Chineseness: Legacies of Chinese commercial hegemony in the Cambodian silk industry. *Journal of Contemporary Asia*, 42(2), 210-229. <https://doi.org/10.1080/00472336.2012.668349>
- Davis, Erik W. (2009). *Treasures of the Buddha: Imagining death and life in contemporary Cambodia* (Publication Number 3362007) [Doctoral dissertation, The University of Chicago]. ProQuest Dissertations Publishing.
- Davis, Erik W. (2013). “Khmer Spirits, Chinese Bodies: Chinese Spirit Mediums and Spirit Possession Rituals in Contemporary Cambodia.”. In T. Reuter & A. Horstmann (Eds.), *Faith in the Future: Understanding the Revitalization of Religions and Cultural Traditions in Asia* (pp. 177-196). Brill. [https://doi.org/10.1163/9789004233669\\_010](https://doi.org/10.1163/9789004233669_010)
- Dean, William. (1841). *First lessons in the Tie-chiw dialect*. Siam.
- Dixon, R. M. W. (1979). Ergativity. *Language*, 55, 59-138. <https://doi.org/10.2307/412519>
- Dobson, W. A. C. H. (1962). *Early archaic Chinese: a descriptive grammar*. University of Toronto Press.
- Dong, Hongyuan. (2014). *A history of the Chinese language*. Routledge.
- Dorais, Louis-Jacques. (1991). Refugee adaptation and community structure: The Indochinese in Quebec City, Canada. *International Migration Review*, 25(3), 551-573. <https://doi.org/10.1177/019791839102500305>
- Dorian, Nancy C. (1981). Language change in dying ESG: Fluent speakers’ Gaelic and semi-speakers’ Gaelic. In *Language Death* (pp. 114-156). University of Pennsylvania Press. <https://doi.org/10.9783/9781512815580-008>
- Dryer, Matthew S. (2013). Order of relative clause and noun. In M. S. H. Dryer, Martin (Ed.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology.

- Dryer, Matthew S. (2014). Competing methods for uncovering linguistic diversity: The case of definite and indefinite articles (Commentary on Davis, Gillon, and Matthewson). *Language*, 90(4), e232-e249. <https://doi.org/10.1353/lan.2014.0070>
- Duffus, William. (1883). *English-Chinese vocabulary of the vernacular or spoken language of Swatow*. English Presbyterian Mission Press.
- Edwards, Penny. (2002). Time travels: Locating *xinyimin* in Sino-Cambodian histories. In P. Nyíri & I. Saveliev (Eds.), *Globalizing Chinese Migration* (pp. 254-289). Ashgate.
- Edwards, Penny. (2008). Between a song and a *prei*: Tracking Cambodian history and cosmology through the forest. In Anne Ruth Hansen & Judy Ledgerwood (Eds.), *At the edge of the forest: Essays on Cambodia, history, and narrative in honor of David Chandler* (pp. 137-162): Cornell University Press.
- Edwards, Penny. (2009). Ethnic Chinese in Cambodia. In *Ethnic Groups in Cambodia* (pp. 174-234). Center for Advanced Study.
- Edwards, Penny. (2012). Addendum to the introduction: New directions. *Cross-Currents: East Asian History and Culture Review*(4), 8-11. <https://escholarship.org/uc/item/1nn699nv>
- Edwards, Penny. (2012b). Endnote to "Mediating Chineseness in Cambodia" - Sojourns across sources: Unbraiding Sino-Cambodian histories. *Cross-Currents: East Asian History and Culture Review*(4), 118-136. <https://escholarship.org/uc/item/0bs2358q>
- Egerod, Søren. (1958). Swatow loan words in Siamese. *Acta Orientalia*, 23, 137-156. <https://doi.org/10.5617/ao.5303>
- Eiampailin, Janida. (2004). *The phonological interference of Swatow in Standard Thai by Chinese speakers in Bangkok* [Master's thesis, Mahidol University].
- Erbaugh, Mary S. (2002). Classifiers are for specification: Complementary functions for sortal and general classifiers in Cantonese and Mandarin. *Cahiers de linguistique Asie orientale*, 31(1), 33-69.

- Fielde, Adele Marion. (1878). *First lessons in the Swatow dialect*. Swatow Printing Office Company.
- Fielde, Adele Marion. (1883). *A pronouncing and defining dictionary of the Swatow dialect: Arranged according to syllables and tones*. American Presbyterian Mission Press.
- Filippi, Jean-Michel. (2010, November 26). Cambodia: The swing of the pendulum. *Phnom Penh Post*. <https://www.phnompenhpost.com/post-plus/cambodia-swing-pendulum>
- Filippi, Jean-Michel. (2013, February 8). A history of the Chinese in Cambodia. *Phnom Penh Post*. <https://www.phnompenhpost.com/post-plus/history-chinese-cambodia>
- Filippi, Jean-Michel, & Hiep, Chan Vicheth. (2016). *Khmer pronouncing dictionary: Standard Khmer and Phnom Penh dialect*. Phnom Penh: KAM, UNESCO.
- Fishman, Joshua A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Multilingual Matters.
- Francis, Elaine J., & Matthews, Stephen. (2005). A multi-dimensional approach to the category ‘verb’ in Cantonese. *Journal of Linguistics*, 41(2), 269-305.
- Garellek, Marc. (2013). *Production and perception of glottal stops* [Dissertation, University of California, Los Angeles]. <https://escholarship.org/uc/item/7zk830cm>
- Giles, Herbert Allen. (1877). *Handbook of the Swatow dialect: With a vocabulary*. Published with the assistance of the Straits’ Government.
- Givón, Talmy. (1983). Topic continuity in discourse: an introduction. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative cross-language study* (pp. 1-41). <https://doi.org/10.1075/tsl.3.01giv>
- Goddard, Josiah. (1888). *A Chinese and English vocabulary, in the Tie-chiu dialect* (2 ed.). American Presbyterian Mission Press.
- Goh, Eng Choon (吴英俊). (2017). 潮州口语 (*Colloquial Teochew*). NETUCC.

- Goh, Eng Choon (吴英俊). (2018). 坦坦潮州话 (*Let's speak Teochew*). NETUCC.
- Goh, Eng Choon (吴英俊). (2020). 新编潮州口语集释 (*Colloquial Teochew: New edition*). NETUCC.
- Goh, Rui Long (吴芮珑). (2011). 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]. Digital Repository of NTU.  
<http://hdl.handle.net/10356/43682>
- Gong, Gunhu. (2000). *A list of Swatow loanwords in Thai* [Unpublished doctoral dissertation]. Shantou University.
- Greenberg, Joseph H. (1972). Numeral classifiers and substantival number: Problems in the genesis of a linguistic type. *Working Papers on Language Universals*, 9, 1-39.
- Greenberg, Joseph H. . (1990). Dynamic aspects of word order in the numeral classifier. In K. Denning & S. Kemmer (Eds.), *On language: Selected writings of Joseph H. Greenberg* (pp. 227-240). Stanford University Press.
- Gruzdeva, Ekaterina. (2004). Numeral classifiers in Nivkh. *Sprachtypologie und Universalienforschung (STUF)*, 57(2/3), 300-329.  
<https://doi.org/10.1524/stuf.2004.57.23.300>
- Gruzdeva, Ekaterina. (2015). Sociolinguistic and linguistic outcomes of Nivkh-Russian language contact. In C. Stolz (Ed.), *Language empires in comparative perspective* (pp. 153-182). de Gruyter. <https://doi.org/10.1515/9783110408362.153>
- Gyarunsut, Pranee (ปราณี กาย อรุณ สิทธิ). (1983). คำ ยืม ภาษา จีน ใน ภาษา ไทย ปัจจุบัน (*Chinese loanwords in modern Thai*) [Master's thesis, Chulalongkorn University]. Chulalongkorn University Intellectual Repository.  
<http://cuir.car.chula.ac.th/handle/123456789/37061>

- Haiman, John. (1983). Iconic and economic motivation. *Language*, 59(4), 781-819.  
<https://doi.org/10.2307/413373>
- Handel, Zev. (2015). The classification of Chinese: Sinitic (The Chinese language family). In W. S.-Y. Wang & C. Sun (Eds.), *The Oxford handbook of Chinese linguistics* (pp. 34-44). Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780199856336.013.0001>
- Haspelmath, Martin. (1997). *Indefinite pronouns*. Oxford University Press.  
<https://doi.org/10.1093/oso/9780198235606.001.0001>
- Haspelmath, Martin. (2004). Coordinating constructions: An overview. In M. Haspelmath (Ed.), *Coordinating constructions* (pp. 3-40). John Benjamins.  
<https://doi.org/10.1075/tsl.58.03has>
- Haspelmath, Martin. (2009). Lexical borrowing: Concepts and issues. In Martin Haspelmath & Uri Tadmor (Eds.), *Loanwords in the world's languages: A comparative handbook* (pp. 35-54): De Gruyter Mouton.
- Haspelmath, Martin. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language*, 86(3), 663-687.
- Haspelmath, Martin. (2016). The serial verb construction: Comparative concept and cross-linguistic generalizations. *Language and Linguistics*, 17(3), 291-319.  
<https://doi.org/10.1177/2397002215626895>
- Haspelmath, Martin, & Sims, Andrea D. (2010). *Understanding morphology* (2nd ed.). Hodder.  
<https://doi.org/10.4324/9780203776506>
- Haspelmath, Martin, with Oda Buchholz. (1998). Equative and similitive constructions in the languages of Europe. In Johan van der Auwera (Ed.), *Adverbial constructions in the languages of Europe* (pp. 277-334): Mouton de Gruyter.  
<https://doi.org/10.1515/9783110802610.277>

- Heine, Bernd, & Kuteva, Tania. (2002). *World lexicon of grammaticalization*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511613463>
- Himmelmann, Nikolaus P. (1998). Documentary and descriptive linguistics. *Linguistics*, 36, 161-195.
- Ho, Pei Qin (何佩芬). (2009). 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, Retrieved from <https://hdl.handle.net/10356/14980> Digital Repository of NTU database.
- Hock, Hans Henrich. (1991). *Principles of historical linguistics*. De Gruyter Mouton. <https://doi.org/10.1515/9783110746440>
- Hong, Ying. (2013). *A phonetic study of Chaozhou Chinese* [Doctoral thesis, Hong Kong University of Science and Technology]. HKUST SPD: The Institutional Repository. <http://hdl.handle.net/1783.1/62308>
- Huang, Yu (黄瑜), & Mo, Yuan-Yuan (莫源源). (2017). 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, 95-97.
- Huffman, Franklin Eugene. (1967). *An outline of Cambodian grammar* (Publication Number 6717239) [Doctoral thesis, Cornell University]. ProQuest Dissertations & Theses Global.
- Kang, Yoonjung. (2011). Loanword phonology. In Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, & Keren Rice (Eds.), *The Blackwell companion to phonology* (pp. 1-25): John Wiley & Sons.
- Karlgren, Bernhard. (1915). *Études sur la phonologie chinoise*. Brill.
- Kasanga, Luanga Adrien. (2012). Mapping the linguistic landscape of a commercial neighbourhood in Central Phnom Penh. *Journal of Multilingual and Multicultural Development*, 33(6), 553-567. <https://doi.org/10.1080/01434632.2012.683529>

- Kawachi, Kazuhiro. (2007). *A grammar of Sidaama (Sidamo), a Cushitic language of Ethiopia* (Publication Number 3268817) [Doctoral dissertation, State University of New York at Buffalo]. ProQuest Dissertations & Theses Global.
- Kaysina, Inna. (2015). Grammatical effects of Russian-Udmurt language contact. In C. Stolz (Ed.), *Language empires in comparative perspective* (pp. 219-236). De Gruyter.  
<https://doi.org/10.1515/9783110408362.219>
- Khoo, Kiak Uei (邱克威). (2017a). 雪兰莪丹絨土拔村潮州方言(澄海外砂话)音系调查分析 (Investigation and analysis on the phonology of Teochew dialect (Chengwaisha dialect) in Tanjung Sepat Village, Selangor). In W. Liao (Ed.), *马来西亚华人民俗研究论文集 (Collected essays on Malaysian Chinese folk culture)* (pp. 199-217). Strategic Information and Research Development Centre.
- Khoo, Kiak Uei (邱克威). (2017b). 马来西亚汉语方言声调变异及其社会因素的调查研究: 以霹靂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, 258-267.
- Khoo, Kiak Uei (邱克威). (2017c). 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), 14-24.  
<https://ejournal.um.edu.my/index.php/JCLC/article/view/17465>
- Khoo, Kiak Uei (邱克威). (2018). 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), 201-212.  
[https://doi.org/10.6710/JTLL.201810\\_13\(2\).0002](https://doi.org/10.6710/JTLL.201810_13(2).0002)
- Kiernan, Ben. (1986). Kampuchea's ethnic Chinese under pol pot: A case of systematic social discrimination. *Journal of Contemporary Asia*, 16(1), 18-29.

- Kiernan, Ben. (1990). The survival of Cambodia's ethnic minorities. *Cultural Survival Quarterly*, 14(3), 64-66.
- König, Ekkehard, & Siemund, Peter. (2007). Speech act distinctions in grammar. In T. Shopen (Ed.), *Language typology and syntactic description* (2 ed., Vol. 1, pp. 276-324). Cambridge University Press.
- Koptjevskaja-Tamm, Maria. (2002). Adnominal possession in the European languages: Form and function. *Sprachtypologie und Universalienforschung (STUF)*, 55(2), 141-172.  
<https://doi.org/10.1524/stuf.2002.55.2.141>
- Koptjevskaja-Tamm, Maria. (2003). *A woman of sin, a man of duty, and a hell of a mess: Non-determiner genitives in Swedish*. In F. Plank (Ed.), *Noun phrase structure in the languages of Europe* (pp. 515-558). Mouton de Gruyter.  
<https://doi.org/10.1515/9783110197075.3.515>
- Kwok, Bit-Chee. (2018). *Southern Min: Comparative phonology and subgrouping*. Routledge.
- Kyne, Phelim. (1999, June 25). Cambodia's middle kingdom: Chinese schools, back from the brink. *Phnom Penh Post*. <https://www.phnompenhpost.com/national/chinese-schools-back-brink>
- Ladefoged, Peter, & Disner, Sandra Ferrari. (2012). *Vowels and consonants*. John Wiley & Sons.
- Lambrecht, Knud. (1988). There was a farmer had a dog: Syntactic amalgams revisited. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*,
- Langacker, Ronald W. (1987). *Foundations of cognitive grammar, Volume I: Theoretical prerequisites*. Stanford University Press.
- LaPolla, Randy J. (1990). *Grammatical relations in Chinese: Synchronic and diachronic considerations*. [Doctoral dissertation, University of California, Berkeley].

- LaPolla, Randy J. (1993). Arguments against 'subject' and 'direct object' as viable concepts in Chinese. *Bulletin of the Institute of History and Philology*, 63(4), 759-813.
- LaPolla, Randy J. (1995). Pragmatic relations and word order in Chinese. In P. Downing & M. P. Noonan (Eds.), *Word order in discourse* (pp. 297-329).  
<https://doi.org/10.1075/tsl.30.11lap>
- LaPolla, Randy J. (2006). On grammatical relations as constraints on referent identification. In T. Tsunoda & T. Kageyama (Eds.), *Voice and grammatical relations* (pp. 139-151). John Benjamins Publishing Company. <https://doi.org/10.1075/tsl.65.09lap>
- LaPolla, Randy J. (2010). Language contact and language change in the history of the Sinitic languages. *Procedia-Social and Behavioral Sciences*, 2(5), 6858-6868.
- LaPolla, Randy J. (2017). Noun-modifying clause constructions in Sino-Tibetan languages. In Y. Matsumoto, B. Comrie, & P. Sells (Eds.), *Noun-modifying clause constructions in languages of Eurasia: Reshaping theoretical and geographical boundaries* (pp. 91-103).  
<https://doi.org/10.1075/tsl.116.05lap>
- Lau, Elaine. (2016). *Acquisition of relative clauses in Cantonese: A multi-factorial analysis* [Doctoral dissertation, University of Hawai'i at Manoa].
- Learn Dialect in Singapore. (2020, February 4). *[In Singapore Teochew] COVID-19 (Wuhan coronavirus) and the precautions that you can take [Video]*. YouTube.  
<https://www.youtube.com/watch?v=qAPx9tg42t0>
- Lee, Cher Leng. (2015). 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* (pp. 196-215). Routledge.  
<https://doi.org/10.4324/9781315759371>
- Lee, Cher Leng (李子玲). (2003). 新加坡人说的潮州话 (Teochew spoken by Singaporeans). In C. H. Lee (Ed.), *海外潮人的移民经验 (The migration experience of the overseas Teochew community)* (pp. 240-260). Global Publishing Company.  
<https://doi.org/10.1142/g100>

- Lee, Hugo Yu-Hsiu. (2014). Losing Chinese as the first language in Thailand. *Asian Social Science*, 10(6), 176-193. <https://doi.org/10.5539/ass.v10n6p176>
- Levin, Beth. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago Press.
- Li, Charles N., & Thompson, Sandra A. (1981). *Mandarin Chinese: A functional reference grammar*. University of California Press.
- Li, Danni. (2014). *Acoustic analysis of the tones in the Shantou dialect* [Master's thesis, University of Massachusetts Amherst]. ScholarWorks@UMass Amherst. [https://scholarworks.umass.edu/masters\\_theses\\_2/29](https://scholarworks.umass.edu/masters_theses_2/29)
- Li, Wei, Saravanan, Vanithamani, & Ng, Julia Lee Hoon. (1997). Language shift in the Teochew community in Singapore: A family domain analysis. *Journal of Multilingual and Multicultural Development*, 18(5), 364-384. <https://doi.org/10.1080/01434639708666326>
- Li, Xuping, & Bisang, Walter. (2012). Classifiers in Sinitic languages: From individuation to definiteness-marking. *Lingua*, 122(4), 335-355. <https://doi.org/10.1016/j.lingua.2011.12.002>
- Li, Xinkui (李新魁), & Lin, Lunlun (林伦伦). (1992). 潮汕方言詞考釋 (*Lexical research on Chaoshan dialect*). Guangdong People's Publishing House.
- Li, Yongming (李永明). (1959). 潮州方言 (*Chaozhou dialect*). 中華書局.
- Li, Yongming (李永明). (1986). 潮州方言语音的内部差别 (Phonetic differences among Chaozhou dialects). 湘潭大学学报 (社会科学版)(2), 93-97.
- Li, Yongming (李永明). (1991). 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, 1, 56-63.

- Li, Yiyan (李以严), & Weng, Jingquan (翁敬铨). (1987). 潮汕方言的比较句 (Comparative structures in the Chaoshan dialect). 韩山师专学报: 社会科学版(2), 93-95.
- Lim, Hiong Seng. (1886). *Handbook of the Swatow vernacular*. Koh Yew Hean Press.
- Lim, Alvin Cheng-Hin. (2011). *Desiring Cambodia* [Unpublished doctoral dissertation, University of Hawai'i at Mānoa].
- Lin, Lien Hsien. (1973). 潮州方言比较研究—从潮州方言与广韵的比较论潮语在汉语方言中的地位 (*A comparative study of Ch'ao-chow (Chiu-chaw) dialect with special reference to Kuang-yun*) [Doctoral thesis, University of Hong Kong]. The HKU Scholars Hub. <http://hdl.handle.net/10722/34621>
- Lin, Lunlun (林伦伦). (1991). 潮汕方言实词的几种词法特点 (Several lexical features of content words in Chaoshan dialect). 汕头大学学报: 人文科学版, 7(2), 62-69.
- Lin, Lunlun (林伦伦). (1992). 潮汕方言的虚词及其语法意义 (Function words in Chaoshan dialect and their grammatical meanings). 汕头大学学报: 人文科学版, 8(1), 53-61.
- Lin, Lunlun (林伦伦). (1994). 广东澄海方言音系记略 (A brief description of the sound system of Chenghai vernacular in Eastern Guangdong province). 汕头大学学报: 人文社会科学版(1), 82-91.
- Lin, Lunlun (林伦伦). (1994). 广东揭西县方音研究 (A brief description of the phonetic features found in two vernaculars at Jiexi County in Eastern Guangdong province). 汕头大学学报: 人文社会科学版, 10(3), 82-89.
- Lin, Lunlun (林伦伦). (1995). 潮汕方言声调研究 (Research on the tones of Chaoshan dialect). 语文研究, 1, 52-59.
- Lin, Lunlun (林伦伦). (1996). 澄海方言研究 (*Research on the Chenghai dialect*). 汕头大学出版社.

- Lin, Lunlun (林伦伦). (2006). 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, 24-30.
- Lin, Lunlun (林伦伦). (2015). 潮汕方言历时研究 (*A diachronic study of Chaoshan dialect*). Jinan University Press.
- Lin, Lunlun (林伦伦), & Chen, Xiaofeng (陈小枫). (1996). 广东闽方言语音研究 (*Phonetic research of Guangdong Min dialect*). Shantou University Press.
- Lin, Lunlun (林伦伦), & Chen, Zhaoer (陈照儿). (2011). 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), 56-60.
- Lin, Lunlun (林伦伦), Lin, Chunyu (林春雨), & Xu, Zemin (许泽敏). (2005). 南澳岛闽方言语音记略 (A comparative study of the sounds of the Min dialects in Nan'ao Island). *汕頭大學學報: 人文社會科學版*, 21(2), 73-76.
- Lin, Qing. (2019). *The diachrony of tone sandhi: Evidence from Southern Min Chinese*. Peking University Press; Springer.
- Low, Elissa Jia Min. (2014). *A sketch grammar of Singapore Teochew* [Final year project, Nanyang Technological University]. Digital Repository of NTU. <https://hdl.handle.net/10356/65807>
- Lü, Shuxiang (吕叔湘). (1985). 近代汉语指代词 (*Modern Chinese pronouns*) (Lansheng Jiang (江蓝生) Ed.): 学林出版社.
- Ly, Y. (2000). *Heaven becomes Hell: A survivor's story of life under the Khmer Rouge* (J. S. Driscoll, Ed.). Yale University Southeast Asia Studies.
- Manomaivibool, Prapin. (1975). *A study of Sino-Thai lexical correspondences* [Doctoral dissertation, University of Washington]. ResearchWorks Archive. <http://hdl.handle.net/1773/11132>

- Marks, Paul. (2000). China's Cambodia strategy. *Parameters*, 30(3), 92-108.
- Matisoff, James A. (2003). *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. University of California Press.
- Matthews, Stephen. (1999). YR Chao and universal Chinese grammar. In D. Cram, A. R. Linn, & E. Nowak (Eds.), *History of linguistics 1996, Volume 1: Traditions in linguistics worldwide* (pp. 217-226). John Benjamins Publishing Company.  
<https://doi.org/10.1075/sihols.94.27mat>
- Matthews, Stephen. (2006). On serial verb constructions in Cantonese. In A. Y. Aikhenvald & R. M. W. Dixon (Eds.), *Serial verb constructions: A cross-linguistic typology* (pp. 69-87).
- Matthews, Stephen, & Pacioni, Patrizia. (1997). Specificity and genericity in Cantonese and Mandarin. In L. Xu (Ed.), *The referential properties of Chinese noun phrases* (pp. 45-59). École des Hautes Études en Sciences Sociales, Centre de Recherches Linguistiques sur l'Asie Orientale.
- Matthews, Stephen, Xu, Huiling, & Yip, Virginia. (2005). Passive and unaccusative in the Jieyang dialect of Chaozhou. *Journal of East Asian Linguistics*, 14(4), 267-298.  
<https://www.jstor.org/stable/20100893>
- Matthews, Stephen, & Xu, Hui Ling. (2002). The pronominal copula in Chaozhou. Paper presented at the *11th Annual Conference of the International Association of Chinese Linguistics*, Nagoya, Japan.
- Matthews, Stephen, & Yip, Virginia. (2008). Passive, unaccusative and pretransitive constructions in Chaozhou. In R. Djamouri, B. Meisterernst, & R. Sybesma (Eds.), *Chinese linguistics in Leipzig (漢語語言學在萊比錫)* (pp. 163-174). EHESS, Centre de Recherches Linguistiques sur l'Asie Orientale.

- McFarland, Joanna R. (2017). *A preliminary comparative analysis of Cambodian Teochew: Evidence for contact-induced change* [Unpublished master's dissertation, University of Hong Kong].
- McFarland, Joanna R. (2021). Language contact and lexical changes in Khmer and Teochew in Cambodia and beyond. In T. Hoogervorst & C. Chia (Eds.), *Sinophone Southeast Asia* (pp. 91-128). Brill. [https://doi.org/10.1163/9789004473263\\_005](https://doi.org/10.1163/9789004473263_005)
- McGee, Matthew. (2017). Ethnic Policies toward the Viet-Cambodians and Sino-Cambodians in the People's Republic of Kampuchea, 1979-1989. *Ezra's Archives*, 7(1), 52-75. <https://hdl.handle.net/1813/49729>
- Miao, Ruiqin. (2005). *Loanword adaptation in Mandarin Chinese: Perceptual, phonological and sociolinguistic factors*. (Doctoral dissertation), Stony Brook University, Retrieved from <https://rucore.libraries.rutgers.edu/rutgers-lib/38301/> RUCore: Rutgers University Community Repository database.
- Morita, Liang. (2007). Discussing assimilation and language shift among the Chinese in Thailand. *International journal of the sociology of language*, 2007(186), 43-58. <https://doi.org/10.1515/IJSL.2007.041>
- Myers, James, & Li, Yingshing. (2009). Lexical frequency effects in Taiwan Southern Min syllable contraction. *Journal of Phonetics*, 37(2), 212-230. <https://doi.org/10.1016/j.wocn.2009.02.002>
- Myers-Scotton, Carol. (1993). *Duelling languages: Grammatical structure in codeswitching*. Oxford University Press.
- Ng, Julia Lee Hoon. (1996). *Language maintenance and language shift in the Teochew Chinese community* [Academic exercise, National Institute of Education]. NIE Digital Repository. <http://hdl.handle.net/10497/1723>
- Noonan, Michael. (2007). Complementation. In T. Shopen (Ed.), *Language typology and syntactic description: Complex constructions* (2nd ed., Vol. 2, pp. 52-150).

- Norman, Jerry. (1988). *Chinese*. Cambridge University Press.
- Nyíri, Pál. (2012). 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), 369-397. <https://doi.org/10.1353/ach.2012.0012>
- Nyíri, Pál. (2015). “New migrants” from the PRC and the transformation of Chinese media: The case of Cambodia. In W. Sun & J. Sinclair (Eds.), *Media and communication in the Chinese diaspora* (pp. 15-31). Routledge. <https://doi.org/10.4324/9781315717265>
- Ong, Teresa Wai See. (2018). *Language maintenance in Malaysia: A case study of the Chinese community in Penang* [Doctoral thesis, Griffith University]. Griffith Research Online. <http://hdl.handle.net/10072/382738>
- Pan, Jiayi (潘家懿). (2000). 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)*. 北京语言文化大学出版社 (Beijing Language and Culture University Press).
- Panh, Rithy, & Bataille, Christophe. (2012). *The elimination: A survivor of the Khmer Rouge confronts his past and the commandant of the killing fields* (J. Cullen, Trans.). Other Press.
- Paul, Waltraud. (2008). The *serial verb construction* in Chinese: A tenacious myth and a Gordian knot. *The Linguistic Review*, 25(3-4), 367-411. <https://doi.org/10.1515/TLIR.2008.011>
- Peng, Anne Elise. (2011). Head-final and head-initial relative clauses in Jambi Teochew. In K. Otaki, H. Takeyasu, & S.-i. Tanigawa (Eds.), *Online Proceedings of GLOW in Asia Workshop for Young Scholars* (pp. 262-276). [http://faculty.human.mie-u.ac.jp/~glow\\_mie/Workshop\\_Proceedings/20Peng.pdf](http://faculty.human.mie-u.ac.jp/~glow_mie/Workshop_Proceedings/20Peng.pdf)
- Peng, Anne Elise. (2012). *Aspects of the syntax of Indonesian Teochew* [Unpublished doctoral dissertation, University of Delaware].

- Phadungsrisavas, Veerakit. (2008). *An acoustical comparative study of Swatow Chinese tones spoken in five regions of Thailand* [Unpublished master's thesis, Mahidol University].
- Pou, Saveros, & Jenner, Philip N. (1973). Some Chinese loanwords in Khmer. *Journal of Oriental Studies*, 11(1), 1-90.
- Quek, Huan Ting Geraldine. (2013). *Investigating language loss within Singapore's Teochew community* [Unpublished final year project, Nanyang Technological University].
- Rappa, Antonio. (2014). *Thailand's Chinese population: Teochiu speakers and political identity* [Unpublished paper].
- Ring, Hiram. (2017). praascripts. <https://github.com/lingdoc/praaascripts>
- Romaine, Suzanne. (2010). Contact and language death. In R. Hickey (Ed.), *The handbook of language contact* (pp. 320-339). Blackwell Publishing Ltd.  
<https://doi.org/10.1002/9781444318159.ch16>
- Sak-Humphry, Chhany. (2016). *Colloquial Cambodian: The complete course for beginners (new edition)*. Routledge. <https://doi.org/10.4324/9780203120651>
- San Roque, Lila, Gawne, Lauren, Hoenigman, Darja, Miller, Julia Colleen, Rumsey, Alan, Spronck, Stef, Carroll, Alice, & Evans, Nicholas. (2012). Getting the story straight: Language fieldwork using a narrative problem-solving task. *Language Documentation & Conservation*, 6, 135-174. <http://hdl.handle.net/10125/4504>
- Schliesinger, Joachim. (2011). *Ethnic groups of Cambodia, volume 3: Profile of Austro-Thai and Sinitic-speaking peoples* (Vol. 3). White Lotus.
- Schmidt, Annette. (1983). *Young people's Dyirbal: An example of language death from Australia* [Master's thesis, Australian National University.]. Australia National University Digital Collections. <https://openresearch-repository.anu.edu.au/handle/1885/132710>

- Self, Stephen. (2014). Another look at serial verb constructions in Khmer. *Mon-Khmer Studies*, 43.1, 84-102.
- Selinker, Larry. (1992). *Rediscovering interlanguage*. Routledge.
- Shi, Qisheng (施其生). (1996). 方言论稿 (*Collected papers on dialect research*). Guangdong People's Publishing House.
- Shi, Qisheng (施其生). (1999). 汕头方言的代词 (Pronouns in the Shantou dialect). In R. Li & S. Zhang (Eds.), 代词 (*Pronouns*) (pp. 289-324). Jinan University Press.
- Silverstein, Michael. (1976). Hierarchy of features and ergativity. In R. M. W. Dixon (Ed.), *Grammatical categories in Australian languages* (pp. 112-171). Australian Institute of Aboriginal Studies. <https://doi.org/10.5281/zenodo.4688088>
- Sim, Tze Wei. (2012). Why are the native languages of the Chinese Malaysians in decline. *Journal of Taiwanese Vernacular*, 4(1), 62-95.
- Simpson, Andrew, Soh, Hooi Ling, & Nomoto, Hiroki. (2011). Bare classifiers and definiteness: A cross-linguistic investigation. *Studies in Language*, 35(1), 168-193. <https://doi.org/10.1075/sl.35.1.10sim>
- Siphat, Touch. (2017). Chinese capitalism in Cambodian sociopolitical contexts: The role of ethnic Chinese in the Cambodian economy. In Y. Santasombat (Ed.), *Chinese capitalism in Southeast Asia* (pp. 181-205). Palgrave Macmillan. <https://doi.org/10.1007/978-981-10-4696-4>
- Stassen, Leon. (1997). *Intransitive predication*. Oxford University Press.
- Stassen, Leon. (2009). *Predicative possession*. Oxford University Press.
- Stenberg, Josh. (2015). Multilingualism and the West Kalimantan Hakka. In W. Li (Ed.), *Multilingualism in the Chinese diaspora worldwide: Transnational connections and local social realities* (pp. 123-140). Routledge. <https://doi.org/10.4324/9781315759371>

- Stevens, Christine Audrey. (1990). *"New life in the freedom country": Young Cambodians in Adelaide* [Doctoral thesis, University of Adelaide]. Adelaide Research & Scholarship. <http://hdl.handle.net/2440/19370>
- Sun, Yenan. (2022). *Incompleteness under Discussion* (Publication Number 28967364) [Doctoral dissertation, The University of Chicago]. ProQuest Dissertations & Theses Global.
- Swadesh, Morris. (1955). Towards greater accuracy in lexicostatistic dating. *International journal of American linguistics*, 21(2), 121-137. <https://doi.org/10.1086/464321>
- Sweetser, Eve. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511620904>
- Tadmor, Uri. (2009). Loanwords in the world's languages: Findings and results. In Martin Haspelmath & Uri Tadmor (Eds.), *Loanwords in the world's languages: A comparative handbook* (pp. 55-75): De Gruyter Mouton.
- Talmy, Leonard. (1972). *Semantic structures in English and Atsugewi* [Doctoral dissertation, University of California, Berkeley]. <https://escholarship.org/uc/item/5g15p348>
- Tan, Danielle. (2006). *La diaspora chinoise du Cambodge: Histoire d'une identité recomposée* [Unpublished master's thesis, Institut d'Etudes Politiques de Paris].
- TCKnow LLC. (2015). *WhatTCSay* (Version 1.30) [Mobile app]. App Store. <https://apps.apple.com/us/app/whattcsay/id550890802>
- Tea, Ammie Sive. (2018). *Webs of ruin and bloom: An analysis of the overseas Chinese networks in Cambodia* [Bachelor thesis, Wesleyan University]. Wesleyan Library's Digital Library. <https://doi.org/10.14418/wes01.1.1485>
- Tea, Van, & Nov, Sokmady. (2009). The ethnic Chinese in Cambodia: Social integration and renaissance of identity. In *Ethnic Groups in Cambodia* (pp. 235-280). Center for Advanced Study.

- Teo, Kok. (1993). *A sociolinguistic description of the Peranakan Chinese of Kelantan, Malaysia*. (Doctoral dissertation), University of California, Berkeley, Retrieved from <https://escholarship.org/uc/item/2hd9m92j> eScholarship database.
- ter Horst, John. (2008). *Weaving into Cambodia: Trade and identity politics in the (post)-colonial Cambodian silk weaving industry* [Doctoral dissertation, Vrije Universiteit Amsterdam]. VU Research Portal.  
<https://research.vu.nl/files/42179835/complete%20dissertation.pdf>
- Thurgood, Graham. (2003). A subgrouping of the Sino-Tibetan languages: The interaction between language contact, change, and inheritance. In G. Thurgood & R. J. LaPolla (Eds.), *The Sino-Tibetan Languages* (pp. 3-21).
- Thurgood, Graham, & LaPolla, Randy J. (2003). *The Sino-Tibetan languages*. Routledge.
- Van der Auwera, Johan, Dobrushina, Nina, & Goussev, Valentin. (2003). A semantic map for imperative-hortatives. In D. Willems, B. Defrancq, T. Coleman, & D. Noël (Eds.), *Contrastive analysis in language: Identifying linguistic units of comparison* (pp. 44-66). Palgrave Macmillan. [https://doi.org/10.1057/9780230524637\\_3](https://doi.org/10.1057/9780230524637_3)
- Veniranda, Yohana. (2015). *Perfective aspect and negation in Pontianak Teochew* [Doctoral dissertation, University of Delaware]. University of Delaware Library Institutional Repository. <http://udspace.udel.edu/handle/19716/17687>
- Veniranda, Yohana. (2016). Oral and nasal vowels in Pontianak Teochew. *LLT Journal: A Journal on Language and Language Teaching*, 18(2), 107-124.  
<https://doi.org/10.24071/llt.2015.180204>
- Verver, Michiel. (2012). Templates of "Chineseness" and trajectories of Cambodian Chinese entrepreneurship in Phnom Penh. *Cross-Currents: East Asian History and Culture Review*, 1(2), 291-322. <https://doi.org/10.1353/ach.2012.0017>

- Verver, Michiel, & Koning, Juliette. (2018). Toward a kinship perspective on entrepreneurship. *Entrepreneurship Theory and Practice*, 42(4), 631-666.  
<https://doi.org/10.1177/1042258718783431>
- Vihman, Marilyn, & Croft, William. (2007). Phonological development: Toward a “radical” templatic phonology. *Linguistics*, 45(4), 683-725.  
<https://doi.org/10.1515/LING.2007.021>
- Wang Jian. (2015). Bare classifier phrases in Sinitic languages: A typological perspective. In H. M. Chappell (Ed.), *Diversity in Sinitic languages* (pp. 110-133). Oxford University Press.
- Wang, Xiaomei. (2012). *Mandarin spread in Malaysia*. University of Malaya Press.
- Wang, Xiaomei. (2016). The Chinese language in the Asian diaspora: A Malaysian experience. In G. Leitner, A. Hashim, & H.-G. Wolf (Eds.), *Communicating with Asia: The future of English as a global language* (pp. 205-215). Cambridge University Press.  
<https://doi.org/10.1017/CBO9781107477186.014>
- Weinreich, Uriel. (1979). *Languages in contact: Findings and problems*. De Gruyter Mouton.
- Wetzer, Harrie. (1992). “Nouny” and “verby” adjectivals: A typology of predicative adjectival constructions. In M. Kefer & J. v. d. Auwera (Eds.), *Meaning and grammar: Cross-linguistic perspectives* (pp. 223-262). De Gruyter Mouton.  
<https://doi.org/10.1515/9783110851656.223>
- Wierzbicka, Anna. (1980). Coordination: The semantics of syntactic constructions. In *Lingua mentalis: the semantics of natural language* (pp. 223-287). Academic Press.
- Willmott, William E. (1967). *The Chinese in Cambodia*. University of British Columbia Press.
- Willmott, William E. (2006). Cambodia. In L. Pan (Ed.), *The encyclopedia of the Chinese overseas* (2nd ed., pp. 144-150). Editions Didier Millet.

- Wright, Martha Susan. (1983). *A metrical approach to tone sandhi in Chinese dialects* [Doctoral dissertation, University of Massachusetts Amherst]. ScholarWorks@UMass Amherst. <https://scholarworks.umass.edu/dissertations/AAI8310348>
- Wu, Fang (吴芳). (2009). 粤东闽语-n,-ŋ 韵尾的方言地理类型研究 (*Dialects geography and typology research on the nasal endings "-n" & "-ŋ" in Min dialect in east Guangdong province*) [Doctoral dissertation, Jinan University].
- Xie, Runzi (谢润姿). (2008). 广东揭阳方言量词初探 (A Preliminary Study of Measure Words in Jieyang Dialect of Guangdong). 广西教育学院学报 (*Journal of Guangxi Institute of Education*), 5, 130-132.
- Xu, Huiling. (2007). *Aspect of Chaozhou grammar: A synchronic description of the Jieyang variety*. Journal of Chinese Linguistics Monograph Series 22. <https://www.jstor.org/stable/23826160>
- Xu, Huiling, & Matthews, Stephen. (2007). 從動詞到子句標記: 潮州方言和台灣閩南話動詞‘說’和‘看’的虛化過程 (From verb to complementizer: The grammaticalization process for the verbs ‘say’ and ‘see’ in the Chaozhou dialect and Taiwanese Southern Min). *Zhōngguó Yǔwén Yánjiū* 中國語文研究, 23, 61-72.
- Xu, Huiling, & Matthews, Stephen. (2011). On the polyfunctionality and grammaticalization of the morpheme *kai* in the Chaozhou dialect. In F. H. Yap, K. Grunow-Hårsta, & J. Wrona (Eds.), *Nominalization in Asian Languages: Diachronic and typological perspectives* (pp. 109-124). John Benjamins Publishing Company. <https://doi.org/10.1075/tsl.96.03xu>
- Xu, Huiling, & Matthews, Stephen. (2013). 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* (pp. 232-244). Huacheng Publishing House.
- Xu, Hui Ling. (2005). Causal negation in the Jieyang dialect: Interaction of negators with verb types, temporal reference, aspect and modality. In D. Adams (Ed.), *Interventions, interactions & interrelations: School of Languages postgraduate research papers on*

- language and literature* (pp. 175-207). School of Languages, the University of Melbourne.
- Xu, Yuhang (徐宇航). (2009). 潮州方言咸深二攝字音韻尾變化研究 (*On the coda changes of the Xian/Shen rhyme groups in Chaozhou dialects*) [Master's thesis, Chinese University of Hong Kong]. Chinese University of Hong Kong (CUHK) Digital Repository. <https://repository.lib.cuhk.edu.hk/en/item/cuhk-326978>
- Xu, Yuhang (徐宇航). (2013). 十九世紀的潮州方言音系 (The phonological system of the Chaozhou dialect in the nineteenth century). *中國文化研究所學報 (Journal of Chinese Studies)*, 57, 223-244.
- Xu, Yuhang (徐宇航). (2016). 潮州方言百餘年來韻母演變的研究 (Phonological changes of vowels in Chaozhou dialect: From mid-19th to early 21st century). *語言學論叢*, 53, 241-260.
- Yap, Foong Ha, & Iwasaki, Shoichi. (2003). From causatives to passives: A passage in some East and Southeast Asian languages. In E. H. Casad & G. B. Palmer (Eds.), *Cognitive linguistics and non-Indo-European languages* (pp. 419-445). De Gruyter Mouton. <https://doi.org/10.1515/9783110197150.10.419>
- Yap, Foong Ha, & Matthews, Stephen. (2008). The development of nominalizers in East Asian and Tibeto-Burman languages. In M. J. López-Couso & E. Seoane (Eds.), *Rethinking grammaticalization: New perspectives, Typological studies in language 76* (pp. 309-341). John Benjamins. <https://doi.org/10.1075/tsl.76>
- Yeo, Pamela Yu Hui. (2011). *A sketch grammar of Singapore Teochew* [Final year project, Nanyang Technological University]. Digital Repository of NTU. <https://hdl.handle.net/10356/93949>
- Yuan, Jiahua (袁家驊). (1983). *汉语方言概要 (Outline of Chinese dialects)*. Wenzhi Gaige Press.
- Yue-Hashimoto, Anne. (1969). The verb 'to be' in Modern Chinese. In J. W. M. Verhaar (Ed.), *The verb 'be' and its synonyms* (pp. 72-111). D. Reidel Publishing Company.

- Yue-Hashimoto, Anne. (1993). *Comparative Chinese dialectal grammar: Handbook for investigators*. École des Hautes Études en Sciences Sociales, Centre de Recherches Linguistiques sur l'Asie Orientale.
- Zeng, Danlin (曾丹林). (2003). 新加坡潮州话 和中国汕头话: 的词汇比较 (Singapore Teochew and Shantou Teochew: Vocabulary comparison). In C. H. Lee (Ed.), 海外潮人的移民经验 (*The migration experience of the overseas Teochew community*) (pp. 261-284). Global Publishing Company. <https://doi.org/10.1142/g100>
- Zhang, Niina Ning. (2013). *Classifier structures in Mandarin Chinese*. De Gruyter Mouton. <https://doi.org/10.1515/9783110304992>
- Zhang, Shiliang. (2015). *The Wuhan dialect: a hybrid Southwestern Mandarin variety of Sinitic* [Doctoral dissertation, University of Hong Kong]. The HKU Scholars Hub. <http://hdl.handle.net/10722/211145>
- Zhang, Shengyu (张盛裕). (1979). 潮阳方言的重叠式 (Reduplication in Chaoyang). 中国语文 (2), 106-114.
- Zhang, Shengyu (张盛裕). (1979). 潮阳方言的连读变调 (Tone sandhi in Chaoyang). 方言(2), 93-121.
- Zhang, Shengyu (张盛裕). (1980). 潮阳方言的连读变调 (二) (Tone sandhi in Chaoyang (2)). 方言(2), 123-136.
- Zhang, Shengyu (张盛裕). (1981). 潮阳方言的语音系统 (The phonetic system of the Chaoyang dialect). 方言(1), 27-39.
- Zhang, Shengyu (张盛裕). (1982). 潮阳方言的象声字重叠式 (Onomatopoeia in Chaoyang). 方言(3), 181-182.
- Zhang, Shuangqing (张双庆). (2013). 研究海外汉语方言外来语的一些思考 (Some thoughts on studying loanwords in overseas Chinese dialects). 北方语言论丛, 3, 55-63.

- Zhang, Xiaoshan (张晓山). (1992). 潮州话连读 变调的特点 (The characteristics of Teochew). In D. Liang, L. Lin, & Y. Zhu (Eds.), 第二届闽方言学术研讨会论文集 (*Proceedings of the 2nd Min dialect academic conference*) (pp. 199-206). Jinan University Press.
- Zhang, Xiaoshan (张晓山). (1996). 潮州话的否定词“(不会)” (Negative word ‘boi’ in Chaozhou). 韩山师范学院学报 (*Journal of Hanshan Teachers College*)(1), 89-93.
- Zhang, Xiaoshan (张晓山). (1999). 潮州话的否定词“未” (The negative word ‘bue’ in Chaozhou). 韩山师范学院学报 (*Journal of Hanshan Teachers College*)(1), 118-122.
- Zhou, Dagan. (2007). *A record of Cambodia: The land and its people* (P. Harris, Trans.). Silkworm Books.
- Zhuang, Yiyu (庄义友). (2001). 潮州话的否定副词 (Negative adverbs in Chaozhou). 语文研究(3), 47-50.
- Zou, Xun (邹珣). (2007). 澄海方言咸深二摄闭口韵尾的地理分布研究 (*A geographical study of the finals with consonants as [-m]/[-p] to the vocabulary in Xian/Shen rime groups in Chenghai dialect*) [Master's thesis, Jinan University].