Dapeng Dialect:

An Undocumented Cantonese-Hakka Mixed Language in Southern China

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy
in the Graduate School of The Ohio State University

By

Litong Chen, M.A.

Graduate Program in East Asian Languages and Literatures

The Ohio State University

2016

Dissertation Committee:

Dr. Marjorie K.M. Chan, Advisor

Dr. Donald C. Winford

Dr. Zhiguo Xie

Abstract

The Dapeng dialect is a small local dialect spoken by 3,000 to 5,500 speakers in the Dapeng area, Shenzhen, Guangdong Province, China. It is a variety derived from a mixture of Hakka and Cantonese, two of the major varieties of Chinese in Southern China. The Dapeng dialect has hitherto received very little attention from Chinese dialectologists and is still under-documented and insufficiently studied.

This dissertation is built upon both historical records and first-hand fieldwork data collected in the Dapeng area. It takes the initial step towards an extensive collecting of dialect data and a preliminary analysis of the Dapeng dialect and its usage in the Dapeng community. This dissertation is driven by three particular research goals: 1) conducting a detailed description of the contemporary Dapeng dialect, 2) proposing an account of the formation of the dialect, and 3) assessing the vitality of the Dapeng dialect in its speech community.

To achieve the first goal, this dissertation follows the well-established convention of Chinese dialect description, the "dialect report." While describing the Dapeng dialect, this dissertation also makes frequent reference to Standard Chinese, Middle Chinese, Cantonese, and/or Hakka. Results show resemblance between the Dapeng dialect and the source dialects—both Cantonese and Hakka—and the resemblance to the source dialects pertains to all three major linguistic structures: phonology, lexicon, and syntax.

Compared with the Dapeng phonology, which presents a complex hybrid of the two input dialects, the Dapeng lexicon and syntax reflect slightly more similarity to Cantonese.

The second research goal is addressed based on the detailed description of the Dapeng dialect. This dissertation demonstrates that Trudgill's (1986) model of "koineization" is best able to account for the formation of the contemporary Dapeng dialect as induced by the Hakka-Cantonese contact. In particular, levelling and simplification are the two main linguistic processes that gave rise to the present-day Dapeng dialect. This proposal is supported by both linguistic and socio-historical evidence, the latter involving demographic changes in the history of Dapeng, especially with respect to migration history.

In response to the third research goal, this dissertation examines the Dapeng community and assesses the vitality of the Dapeng dialect. After a careful review of the evaluative frameworks, the UNESCO Language Vitality and Endangerment scale is chosen as being the most applicable in the Chinese context. The evaluation is supported by evidence drawn from interviews, observations, and demographic data. The results of the assessment show that the overall vitality of the Dapeng dialect, although only spoken by a small population, is in fact surprisingly positive. The vigorous, healthy condition of the Dapeng dialect is in sharp contrast with many other small Chinese dialects, which are usually reported as being in danger of extinction.

Dedicated to my wife, Dinglei Huang

"*Many women do noble things, but you surpass them all*." (Proverbs 31: 29)

Acknowledgments

I owe my gratitude to all those people who made this dissertation possible.

My heartfelt gratitude first goes to my advisor, Dr. Marjorie K. M. Chan (a.k.a. "Ma Laoshi"). Over the past six years I have learned tremendously both from her extensive knowledge of Chinese linguistics and from her enthusiasm and meticulousness in research. I would have been lost many times without her continuous guidance and encouragement. To me Ma Laoshi is not only an excellent academic advisor but also a mentor and a model in teaching, student service, professionalism, and collegiality. I could never thank her enough.

I am also indebted to Dr. Donald C. Winford. I have benefited enormously from his expertise in contact linguistics and sociolinguistics, ever since my first semester in graduate school at The Ohio State University (OSU). His inspiring suggestions and insightful critiques are priceless in helping to guide me at various stages of this dissertation. It has been my great fortune to have such a patient and knowledgeable linguist in my dissertation committee.

I also owe my gratitude to Dr. Zhiguo Xie. He has been extremely kind and supportive over the years that we have worked together. His careful, thought-provoking comments have greatly helped to focus and shape my ideas, especially in the field of bilingualism and language vitality. I would like to thank him for strictly and consistently holding me to a high research standard.

I want to thank College of Arts and Sciences at OSU for awarding me the University Fellowship (2010-2011) and thank the Department of East Asian Languages

Vita

2008—2009 ....Chinese Teaching Assistant, Department of

Modern and Classical Languages, Allegheny

College

2010...B.A. Chinese Linguistics, Renmin

University of China

2010—2011 ....University Fellow, Department of East

Asian Languages and Literatures, The Ohio

State University

2011—present ....Graduate Teaching Associate of Chinese,

Department of East Asian Languages and

Literatures, The Ohio State University

2012...M.A. Chinese Linguistics, The Ohio State

University

Summer 2015...Lecturer of Chinese, Department of East

Asian Languages and Literatures, The Ohio

State University

Publications

1. Chen, Litong and Tsz-Him Tsui. 2015. Chongqing Mandarin as a contour tone-only language. *Proceedings of Chicago Linguistic Society (CLS-49)* 49:61-68.

2. Chen, Litong. 2015. Variations of the third-person singular pronoun in Hong Kong Cantonese. *University of Pennsylvania Working Papers in Linguistics* 21.1.8:1-5.

3. Chen, Litong and Nicholas Joch. 2015. Cantonese love songs (Yue-ou 粵謳): Preface two. *Asian Literature and Translation* 3(2):1-15.

4. Chen, Litong. 2013. Motivation for code-switching in the Chinese Christian church in the United States. *Texas Linguistic Forum* 56:1-11.

5. Chen, Litong. 2011. The imposition of Cantonese on Mandarin in the city of Guangzhou. *Proceedings of the 23rd North American Conference on Chinese Linguistics (NACCL-23)* 2:93-104.

Fields of Study

Major Field: East Asian Languages and Literatures

Concentration: Chinese Linguistics

Graduate Interdisciplinary Specialization in the Language Sciences

(Minor Area of Focus: Language and Society)

Table of Contents

List of Tables

List of Figures

List of Maps

Chapter 1   Introduction

The Dapeng dialect 大鵬話 is a small local dialect spoken in Dapeng, a district located on the Dapeng Peninsula, Shenzhen, Guangdong Province, in Southern China. The geographic location of Dapeng is shown in Map 1. According to Qiu's (2005) preliminary study, the Dapeng dialect is a mixture of Hakka 客家 (a.k.a. Kejia) and Cantonese 粵 (a.k.a. Yue), two of the major varieties of Chinese in Southern China. The birth and development of the Dapeng dialect are closely related to the history of the Dapeng fortress.

Dapeng was built in 1394 as one of many military fortresses along the South China Sea.[1] It served as a stronghold against frequent pirate raids and foreign invaders (Baoan County Annals Committee 1997, Shenzhen Bowuguan 1997, Ji 2001, Zhang 2006, etc.). The early soldiers and their family members garrisoned there in the late 14th century spoke different Chinese dialects of Guangdong, probably some early variation of Cantonese or Hakka dialects. Due to frequent interaction among these dialect speakers, the early settlers created a "common language" to facilitate communication within and outside the Dapeng fortress. Since Cantonese and Hakka are very different and mostly unintelligible, some kind of "common language" for communication was necessary.

---

[1] The Dapeng fortress is called "大鵬所城 Dàpéng Suǒchéng" in Chinese.

About three centuries later, all civilians in the Dapeng area were evacuated from their land in 1661 (Baoan County Annals Committee 1997, Shenzhen Bowuguan 1997, Guangdong Province Annals Committee 1999, etc.). They were forced to move inland for tens of miles when the central (Qing) government changed the entire Southeastern coast to a restricted military zone.
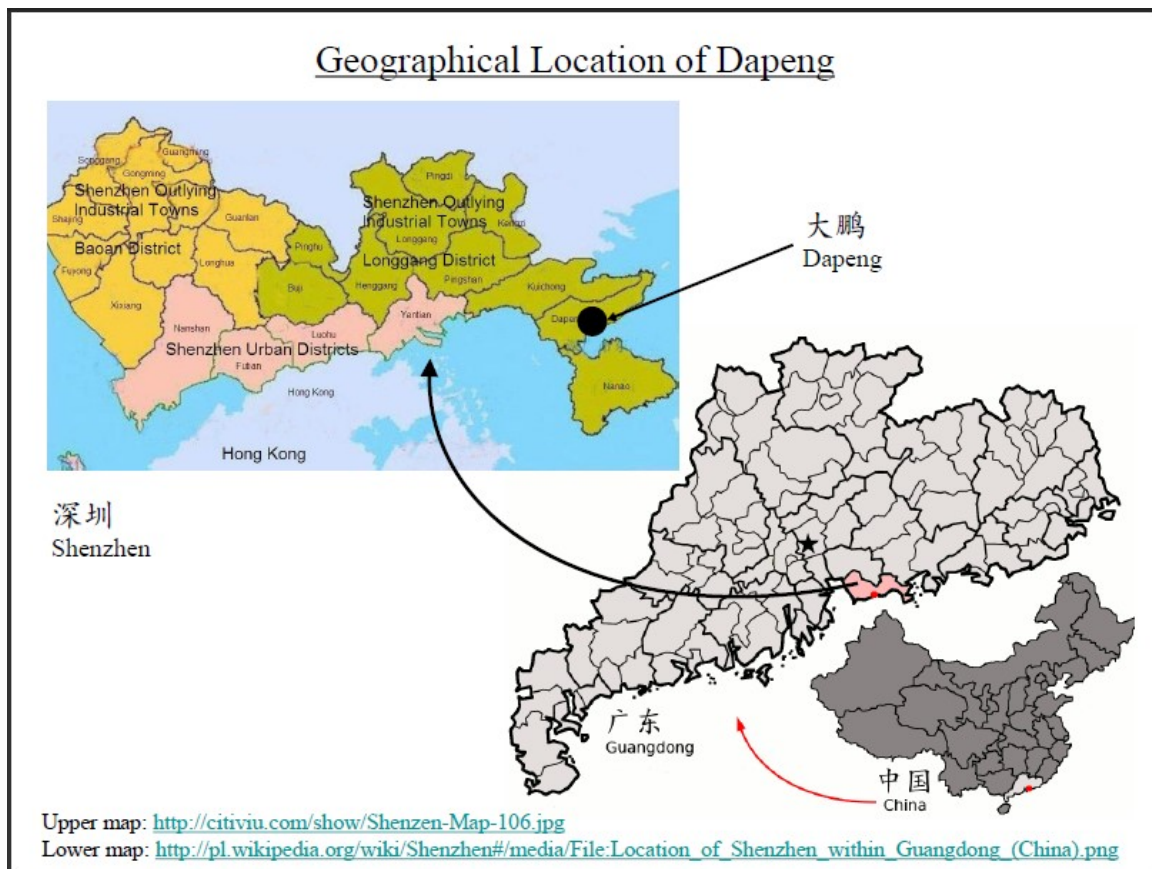
Map 1. Geographical Location of the Dapeng Area[2]

Approximately a decade after the evacuation, in 1670, hundreds of Hakka people from northern and eastern counties of Guangdong were allowed to resettle in the Dapeng peninsula area. By 1688, the total number of Hakka population reached approximately 700 (Cao 1997, Tan 2010, Zeng 2011). Civilian immigration continued to flow into the area after that, but no subsequent large scale waves were recorded. In the meantime, the military population was slowly reduced in the 18th and 19th centuries and completely withdrawn from the Dapeng area in 1899 (Yang and Huang 2001: 153-164).[3] The location of the Dapeng fortress is indicated in Map 2 (Zhang 2006: xxiii).[4]

---

[3] See §5.3 for a more detailed introduction of the history of Dapeng and its demographic changes over time.
[4] This military map was first printed in the Xin'an County Annal (新安縣志) in 1819 and reprined in Zhang (2006).

Map 2. A Military Map of the Dapeng Area

According to an unpublished document provided by the local government, by 2014, the population of regular residents in Dapeng was approximately 3,000. Most of the residents are Dapeng dialect speakers. Including both the non-permanent residents who live in Dapeng periodically and the diaspora population who live in oversea Chinese communities, the total population of Dapeng speakers is about 5,500.[5]

---

[5] This document was obtained from a local government officer during fieldwork conducted in the summer of 2014. This fieldwork is to be introduced in more detail in §2.4, and a more detailed demographic description of current-day Dapeng is given in §6.4.1.

Hakka and Cantonese dialects have been in close contact with each other for centuries in Shenzhen. Hakka dialects are widely spoken in vast areas of Northern and Eastern Guangdong, while Cantonese dialects are prevalent in Western Guangdong as well as in the central part of the province (i.e. the Pearl River Delta). Shenzhen sits right on the watershed that separates these two major dialect groups (Wurn 1987, Baoan County Annals Committee 1997, Zhang 2007). Thus, the hybrid nature of the Dapeng dialect is most likely a result of the constant contact between Hakka and Cantonese.

Another consequence of such contact is the multilingualism that is prevalent in the Dapeng area: Most of the native speakers of the Dapeng dialect are also fluent in Cantonese (for instance, Guangzhou Cantonese, the provincial *lingua franca*) and they are also able to converse in Hakka to a certain degree. In addition to that, in today's China, many Dapeng speakers are also fairly capable in speaking Putonghua (Standard Mandarin, the official language of China).

Due to both the geographic separation of the Dapeng Peninsula from the outside areas of Shenzhen and to the hilly terrain on the peninsula itself, communication between Dapeng and inland towns has always been difficult throughout its history. As a result, today, except for various diaspora communities,[6] the Dapeng dialect is not spoken in any other place outside the peninsula.

---

[6] For instance, New York City has the Tai Pun Residents Association (大鵬同鄉會 *dàpéng tóngxiānghuì*), located in the Manhattan old China Town.

1.1   Research Topics and Questions

A careful review of the literature of Chinese linguistics shows that the Dapeng dialect has hitherto received very little attention from Chinese dialectologists and is still under-documented and insufficiently studied. Apart from Qiu's (2005) brief mention of it in his study of a group of the Army Speech (軍話 *jūnhuà*),[7] dialect islands formed from left over conscripted soldiers from the Ming Dynasty (1368–1644) in Southern China, no other research has been dedicated to the investigation of the hybrid nature of the Dapeng dialect, nor its use in the local community.

This dissertation, built upon first-hand fieldwork data collected in the Dapeng area and upon historical records, therefore takes the initial step towards the study of the Dapeng dialect. In particular, this dissertation aims to address the following three research topics and answer the questions under these topics:

1. Description of the Dapeng dialect

    What exactly is the Dapeng dialect like? What are some of the linguistic features that distinguish it from other Southern Chinese dialects? In precisely what way does this local dialect show its hybrid nature of Hakka and Cantonese? From the perspective of its sound system, lexicon, and syntax, to what degree is it like Hakka? How does it resemble Cantonese?

2. Formation process of the Dapeng dialect

    Based on a detailed description of the mixed nature of the Dapeng dialect, what are some possible formation processes of the Dapeng dialect one could infer? What theoretical framework(s) of dialect formation can be employed to account for the genesis of Dapeng? How well can such

---

[7] Qiu (2005) appears to have conducted some investigation on the Dapeng dialect. But since it does not qualify as Army Speech, the focus of his book, Qiu does not discuss the Dapeng dialect in detail (pages 134-7).

proposal(s) be supported by both linguistic evidence of the dialect and by socio-historical facts?

3. Language vitality of the Dapeng Dialect

Being a local dialect spoken by a small community in Southern China, how much vitality does the Dapeng dialect have in today's peninsula community? Is it endangered, as in the case of many other Chinese local dialects spoken in small communities under the influence of major, more "powerful" Chinese dialects? How do linguistic and social factors (such as bilingualism, language policy, and attitude) affect the maintenance and development of the Dapeng dialect?

All of these questions will be discussed in detail in this dissertation, each corresponding to one or two chapters. The chapters are organized as follows.

1.2   Organization of the Chapters

The current chapter briefly introduces some geographic and sociolinguistic context of the Dapeng dialect. After introducing the organization of the chapters of this dissertation, the remainder of the current chapter outlines some potential contributions of this dissertation to the field of Chinese dialectology.

Chapter 2 offers a critical review of the traditional methodologies of dialect description and fieldwork in the area of Chinese linguistics. This chapter introduces in detail the well-established convention of Chinese dialect description, which is usually referred to in the Chinese dialectology literature as "dialect reports" (方言調查報告 *fāngyán diàochá bàogào*). Comments are also made on its strengths and weaknesses. In addition, Chapter 2 proposes a revised, newer approach to recording and describing the

Dapeng dialect, an approach that aims to take full advantage of the merits of the

conventional methods, while at the same time overcoming its weaknesses. The whole

data collection process of the fieldwork in the Dapeng area is introduced in the chapter,

with the rationale for the design of the fieldwork also explained. In total, I interviewed 20

native speakers of Dapeng from both genders and from different age groups, ranging

from 22 to 84.

Chapters 3 and 4 together provide a detailed description of the Dapeng dialect.

Chapter 3 introduces the main informant and the research procedures. Following the

format of the conventional "dialect report," the chapter then describes the Dapeng

phonetics and its phonological system. References are made to Middle Chinese,

Putonghua, Cantonese, and/or Hakka sound systems, whenever needed. Necessary

changes, albeit minor, are also made in situations when the traditional approach is not

sufficient or accurate enough to describe some particular aspects of the Dapeng sound

system. This chapter relies on the audio recording of Mr. L, one of the senior residents

who, according to both a self-evaluation and peers' assessment, spoke a "standard and

pure" Dapeng dialect.

Chapter 4 also follows the traditional convention of describing the lexicon and

syntax in the format of "dialect report." However, in these sections the description itself

is based on data that were collected via somewhat modified, non-conventional fieldwork

methods, namely, picture naming, acting, or situation explanation. The main reason for

using a newer, more innovative approach was to minimize the interference of Standard

Chinese (both Putonghua, the spoken form, and written Chinese) in the elicitation of

lexical items and grammatical forms. Comparisons among Dapeng, Hakka, and Cantonese will also be made when necessary. Chapter 4 relies on Mr. W, another senior Dapeng speaker, for the lexicon section. The syntax section uses the speaking tasks conducted with all the Dapeng speakers across age groups.

Chapter 5 is built on the language facts presented in the previous descriptive chapters, which have shown that the modern Dapeng dialect is a *result* of the long-term historical Hakka-Cantonese contact. The chapter then proceeds to hypothesize the *processes* (or *mechanisms*)[8] of this contact, which has contributed to the formation of the current-day Dapeng dialect. Specifically, Trudgill's (1986) model of "koineization" is introduced and discussed through a literature review. This chapter demonstrates that the "koineization" model can best account for the formation processes of the contemporary Dapeng dialect, processes that were induced by dialect contact. This proposal is supported by both linguistic and socio-historical evidence, as shown in the first four chapters. Demographic changes in the history of Dapeng are also examined, especially migration history.

Chapter 6 focuses on the contemporary Dapeng dialect community and assesses the language vitality of this local dialect. It starts with a review of some of the most influential frameworks that have been proposed in the literature on language vitality assessment. Comparisons will be made among these models. The UNESCO Language Vitality and Endangerment scale (UNESCO Ad Hoc Expert Group on Endangered

---

[8] Following Winford (2005), the terms "mechanism" and "process" are not distinguished. They are interchangeable here and throughout the entire dissertation.

Languages 2003) is proposed to be the most suitable one, and is used to assess the vitality of the Dapeng dialect. The assessment is supported by interview data, observations, and demographic data. The results of the assessment show that the overall vitality of the Dapeng dialect, although only spoken by a small population, is surprisingly positive, which is in sharp contrast with many other small local dialects in mainland China.

Chapter 7 provides a conclusion for the dissertation and restates the major findings. The chapter also discusses some limitations and ends with suggestions for future research.

1.3   Potential Contributions

This dissertation has the potential to bring both descriptive and theoretical contributions to the field of Chinese dialectology. First, it provides an extensive documentation of the Dapeng dialect, a Hakka-Cantonese mixed local dialect that has barely been studied before. Data were collected and documented by different means: audio-recording, speech-to-text transcription, and observation report. Due to the growing impact from more prestigious dialects, especially from the national standard, Putonghua, many minor colloquial speeches in China are gradually losing their vitality (Cao 2001, Sun 2001, Wu 2008, etc.). Therefore, such documentation is especially crucial for the preservation of a minor, possibly endangered, colloquial speech used by a very small number of speakers like the Dapeng dialect.

This dissertation also provides a detailed, linguistic description of the Dapeng dialect. The description presents some linguistic features that may interest Chinese

dialectologists to further explore this local dialect and other mixed dialects. For instance, the Dapeng sound system shows resemblance with both Cantonese and Hakka, while its lexicon and syntax show more similarity to Cantonese.

Furthermore, since the description is conducted in accordance with the long-practiced convention of Chinese dialectology, it also enables future cross-dialectal comparisons with other dialects recorded in the same framework. This dissertation also sees and discusses both strengths and weaknesses of the convention. In the description of the Dapeng dialect traditional approaches are modified. Such modification also contributes to the improvement of the conventional framework.

From a theoretical perspective, this dissertation brings general theories and models of contact-induced language change to the analyses of mixed Chinese dialect formation, especially a local dialect that contains linguistic elements (such as sounds and vocabulary) from different source dialects (Cantonese and Hakka, in the case of Dapeng). In particular, Trudgill's (1986) model of "koineization" will be discussed and then applied to account for the processes that gave rise to the Dapeng dialect.

Research on koineization of dialects has not been sufficiently conducted in the field of Chinese dialectology. Previous studies on this topic are only a few, and they focus mainly on Northern and Central Chinese dialects (Kuo 2005 and Sun 2012 on Mandarin, Yang 2013 on Wu). In comparison, hybrid dialects as outcomes of Southern Chinese dialect contact have been rarely studied. The current study investigates the formation processes of the Dapeng dialect as induced by the Cantonese-Hakka contact. In this regard, this dissertation could serve as one of the early studies on the topic of mixed

dialect formation in Southern China and hence has the potential of contributing to the knowledge of Chinese contact linguistics in general.

This dissertation also reviews some of the well-established analytical frameworks of language vitality assessment, and attempts to apply these frameworks to the assessment of the vitality of the Dapeng dialect. Based on information collected from first-hand interviews, fieldwork observations, and unpublished demographic data, this dissertation shows the overall positive vitality of the Dapeng dialect and suggests that not all small local dialects in China are necessarily in danger. This finding is another theoretical contribution to Chinese dialectology.

Chapter 2   Methodologies: Description and Fieldwork

The study of modern Chinese dialectology emerged in the early 20th century with the Swedish Sinologist Bernhard Karlgren publishing his masterpiece, *Études sur la phonologie chinoise* ("Studies on Chinese Phonology") between 1915 and 1926. In this pioneering study of Chinese historical phonology, Karlgren presented detailed phonological descriptions of twenty-two Chinese dialects, sixteen being Mandarin dialects, two each from the Wu 吳 and Min 閩 groups, and one each from the Hakka and Yue groups. These materials were collected as he was traveling across China between 1910 and 1912.

Based on the cornerstone laid by Karlgren, a group of Chinese scholars, who were well-trained in general linguistics and in traditional Chinese philology, started to conduct their own fieldwork on Chinese dialects. The descriptive works published afterwards are usually referred to in the Chinese dialectology literature as "dialect reports" (方言調查報告 *fāngyán diàochá bàogào*). A series of early dialect reports published between the 1920's and 1940's provided later descriptive studies of Chinese dialects with good foundational examples.[9] By following and improving on the descriptive methods in these

---

[9] Some of the most frequently cited early works include Yuen Ren Chao's 趙元任 study on Wu dialects in 1928, Changpei Luo's 羅常培 study on Xiamen Min 廈門閩語 in 1931, Xiling Huang's 黃錫凌 study on Guangzhou Yue 廣州粵語 in 1941, and Tonghe Dong's 董同龢 study on Huayang Hakka 華陽客語 in 1948 (Wang 1998: 512-513).

early studies, a conventional approach to describing and comparing Chinese dialects gradually arose.

This chapter starts by introducing the basic format and content of the dialect report (§2.1). Then it offers a critical review of the traditional descriptive methodology of Chinese dialectology in phonetics and phonology as well as in lexicon and syntax (§2.2 and §2.3). Based on this review, this chapter offers revised methods for the data collection and description of a Chinese dialect, which were implemented in the fieldwork trip to the Dapeng area (§2.4).

2.1   Dialect Report

The dialect report is a conventional approach in Chinese dialect description. A dialect report normally includes the following information (Yuan 2001, Li 2007: 9-12, Li and Xiang 2009: 108): 1) introduction of the dialect; 2) background of the informant(s); 3) phonetics and phonology; 4) vocabulary list; 5) syntax; 6) samples of narratives; and 7) maps (optional). Details are given in the remainder of this subsection. [10]

1.   Introduction of the dialect.

This part introduces both the dialect itself and its speech community. Relevant background information often includes historical settings, geographic distribution of the dialect community, demographics (both historical change and current figures), varieties of the dialect, etc.

---

[10] For an example of a detailed descriptive work, see Hashimoto's (1973) report on Meixian Hakka 梅縣客語.

2.   Background of the informant(s).

In this section the investigator provides detailed information about various aspects of the informant(s), including gender, age, family members (esp. when not all his/her immediate relatives speak this dialect as their mother tongue), occupation, education level, linguistic background (esp. language competence of all language/dialects that this person speaks), migration history, etc.

3.   Phonetics and phonology.

To facilitate cross-dialectal comparison, broad phonetic transcription is normally used in this section.

3.1.   Syllable structure (optional). For instance, Hashimoto (1973: 90) describes the Hakka syllable structure as (C)(M)V(E)/T, where C is the initial consonant, M the medial vowel or glide, V the nucleus, E the ending or coda, T the tone. Optional constituents are put in parentheses. This section is not obligatory, since the majority of Chinese dialect reports follow the traditional *initial-final* division of syllables, with initial (聲母 *shēngmǔ*) being C and final (韻母 *yùnmǔ*) containing M, V, and E in Hashimoto's formula. The tone (聲調 *shēngdiào*) is added on to the segment.[11]

3.2.   Sound inventory. This part presents all initial, final, and tone segments in a dialect. Initials are put into a table (聲母表 *shēngmǔ biǎo*) in which the rows distinguish places of articulation and the columns distinguish manners of articulation. Finals are put in a table (韻母表 *yùnmǔ biǎo*) where possible nuclei are listed in different columns, and

---

[11] The syllabic structure of Chinese and the rationale of describing dialect sounds based on syllables will be discussed in more detail in §2.1.

final segments in different rows are distinguished by either the medial or the ending (or both). Tones are distinguished in another table (調目表 *diàomù biǎo*) by their pitch values, and they are usually listed based on their correspondence to the Middle Chinese tonal categories, which then suggest some clues as to the actualization of ancient tones in modern times. Some notes, whenever necessary, are attached to the end of each inventory. These notes often include some clarification and details that are not revealed in the tables alone, such as tone sandhi rules.

3.3.   Syllabary. This part includes combinations of initials, finals, and tones (聲韻調配合表 *shēng yùn diào pèihé biǎo*). When all three sound categories are combined together, theoretically, there can be thousands of possible syllabic combinations in a Chinese dialect. Even if only legitimate ones that exist in the dialect are counted, the total number of combinations may still be in the hundreds. A typical dialect report also includes a syllabary, which lists all combinations of initials, finals, and tones that have emerged from the fieldwork data and thus, this table is also known as "the table of syllables" (音節表 *Yīnjié biǎo*).

3.4.   Morphemes list. Continuing with the organizational logic of the table of syllables, one can sort and organize all of the homophonous morphemes in a dialect, with each syllable recorded during fieldwork. Thus, all the morphemes that share the same phonetic form are grouped together. Given the (roughly) one-to-one correspondence between a morpheme and its orthographic form (i.e. a character), the lengthy, detailed table that sorts out all of the homophonous morphemes according to syllables is also called "the table of homophonous characters" (同音字表 *tóngyīnzì biǎo*).

16

3.5.   Phonological comparison. There is a phonological comparison incorporating both diachronic and synchronic comparisons. The diachronic comparison is usually between the modern dialect and the reconstructed Middle Chinese sound system (11th century), suggesting the historical development of the modern dialect based on patterns of sound correspondence. The synchronic comparison is conducted between the target dialect and *lingua francas*. These *lingua francas* can be some kind of common language on a national level (i.e. Putonghua) or a regional level (for example, Standard Cantonese in Guangdong Province).[12]

Furthermore, synchronic comparison between the target dialect and other related dialects is occasionally seen in the literature, especially in cases where scholars focus on comparative studies among several dialects. This kind of comparison investigates dialects that are either genetically related or typologically similar.[13] In all these three phonological comparisons, again, morphemes and their orthographic forms—Chinese characters—are used as the basis of such comparisons. [14]

---

[12] For instance, in Zhan (2002), the author compares each of the major Yue dialects with Standard Cantonese, which is spoken in Guangzhou; likewise, Chen (1993) makes frequent references to Standard Hakka (that is, Meixian Hakka) when describing the phonological features of Qingxi Hakka 清溪客語, which is spoken in Dongguan, Guangdong.

[13] For instance, in addition to the dialect-*lingua franca* comparison, a large body of Zhan's (2002) work is also devoted to the side-by-side investigation among (non-standard) Yue dialects. Qiu (2005) is a good example of the latter scenario, as he puts a group of Army Speech (軍話 *Jūnhuà*) together, which share no genetic relationship but are strikingly similar in typology.

[14] These comparisons are critical to our understanding of dialect relationships, as Chan (1980: ii) concludes, "against such a backdrop [phonological comparison], it is possible to observe the development of a given dialect with respect not only to earlier strata of the Chinese language, but also to other modern Chinese dialects."

4.   Vocabulary list.

Based on semantic similarities, basic vocabulary items are divided into a number of groups and listed in this section. For each item, the following information is typically provided: pronunciation (in the International Phonetic Alphabet, or "IPA"), orthographic forms (if identifiable), and glosses in Standard Chinese.[15] Some notes may also be added concerning morphological strategies that are particular to this dialect.

5.   Syntax.

This section often includes two parts. One is a summary of syntactic rules, indicating some of the distinctive syntactic features in the dialect.[16] The other part is a set of example sentences, which also contains pronunciation (in IPA), orthographic forms (if identifiable), as well as glosses and translation in Standard Chinese. Since many of the published dialect reports are based on a similar—if not identical—set of example sentences, a syntactic comparison across dialects is made possible.[17]

6.   Samples of narratives.

This part contains transcription of both speaking tasks and spontaneous speeches. This section contains story-telling, picture narration, conversation, and other kinds of spontaneous speech, be it monologic or dialogic. The most famous example might be the

---

[15] For instance, Li and Zhang (1992) include over 900 lexical items in their comparative study of 34 Hakka dialects, covering natural phenomena, animals, vegetables, food, kinship terms, parts of the body, and many other semantic categories.

[16] For instance, Qiu (2005: 47) describes the word order of double object constructions of Pinghai Junhua 平海軍話 as V + Objective$_i$ (direct) + Objective$_j$ (indirect), which is similar to surrounding Yue dialects and is rare among many Chinese dialects.

[17] For instance, Yue-Hashimoto (1993) compiled a handbook for eliciting grammatical structure across Chinese dialects.

story of *The North Wind and the Sun*. Again, pronunciation, orthographic forms (if identifiable), and translation in Standard Chinese are provided in the transcription.

7.   Maps (optional).

Researchers may also attach maps to a dialect report, commonly displaying the geographical location of the dialect speaking community, the distribution of dialect points, and subgrouping schemes.

2.2   Descriptive Methodology: Phonetics and Phonology

The greatest strength of this conventional approach for describing Chinese dialects lies in the way it organizes and presents phonological data from fieldwork. Since phonology has been the focus of dialect investigation since the beginning of the field of Chinese dialectology, the methods of phonological description are well-established. The strengths of the methods are demonstrated in three aspects: 1) it proposes a unified syllable of Chinese; 2) both diachronic comparison with Middle Chinese and synchronic comparison across modern Chinese dialects are emphasized; and 3) the conventional approach investigates doublets and multiple layers of phonological forms. These aspects are discussed in this section. In the meantime, the traditional methods also have their weaknesses. These will also be discussed.

2.2.1   A Unified Syllable Structure of Chinese

First, the conventional approach is applied based on the crucial understanding that modern Chinese dialects have similar typological features. The phonological form of a

morpheme in most Chinese dialects is a single syllable with a tone attached, and the typical Chinese syllable is analyzed as an initial, a final, and a tone. A final can be further divided into a medial, a nucleus, and an ending. Initials are either consonants or semi-vowels; medials are high vowels; nuclei are vowels; and endings are either vowels or nasal/stop consonants. The tone is put over the entire syllable. The entire structure of Chinese is shown below, with the three most important units in bold (viz. initial, final, and tone).

Tone			
Initial	**Final**		
	Medial	Nucleus	Ending
Consonant	*Glide*	*Vowel*	*Glide or nasal/ stop consonant*

Syllable Structure of Chinese

Given the monosyllabic nature of Chinese morphemes, which is by and large universal among Chinese dialects, the conventional approach emphasizes the three critical units—initials, finals, and tones—without involving the details of sub-units (medial, nucleus, and ending). In the context of Chinese dialects in general, it suffices for a dialect report to focus on the syllable level, which directly involves initials, finals, and tones, and not to attempt a description with minute phonetic differences (Yue-Hashimoto 1972: 87).

2.2.2   Diachronic and Synchronic Comparisons

Another strength of the conventional approach, in terms of phonological description, is its clear, strong reference to the Middle Chinese sound system, which is believed to reflect the sound system of 9th and 10th centuries Chinese in the mid to late Tang Dynasty (Norman 1988: 221).[18] Norman and Coblin (1995: 582) point out that while Middle Chinese is not the direct origin of modern Chinese dialects, the phonological categories of the vast majority of modern Chinese dialects have shown a "definite and, by and large, regular relationship" to the reconstructed Middle Chinese sound system.

In Middle Chinese, morphemes, as represented by characters, are assigned to different initial, final, and tonal categories. Since this is also how modern Chinese dialects are described, cross-dialectal comparisons of phonology can be carried out according to these categories in Middle Chinese. In fact, the Middle Chinese sound system functions similarly as a manual to enable dialectologists to systematically examine various parts (initials, finals, and tones) of the modern sound systems to study the correspondence pattern between Middle Chinese and the modern reflexes in a given Chinese dialect.

As a long practiced convention in Chinese dialectology, a dialect report usually includes cross-dialectal comparisons, which rely on reconstructed historical Chinese

---

[18] More precisely, the term "Middle Chinese" used in this dissertation consistently refers to "Late Middle Chinese" as in Pulleyblank's (1984: 3) terminology. Late Middle Chinese is reconstructed based on the rhyme table *Yunjing* 韻鏡 and represents the speech of Chang'an 長安, the capital of the Tang Dynasty, in 9th and 10th centuries. Late Middle Chinese is distinguished from Early Middle Chinese, which is codified in the rhyme dictioinary *Qieyun* 切韻 and reflects the speech of Luoyang 洛陽 in the 6th century (Northern and Southern Dynasties).

sound systems, especially Middle Chinese. Using the initial, final, and tonal categories in the Middle Chinese sound system as a reference, synchronic similarities and differences among modern dialects are indicated by the comparison of diachronic sound changes from Middle Chinese in each dialect. Therefore, sound correspondences between Middle Chinese and modern dialects are especially important for cross-dialectal comparisons. Some features are even regarded as decisive in determining dialectal affiliation.

For instance, if one compares a Wu dialect with a Min dialect, the feature in the former that preserves the three-way distinction of voiceless unaspirated - voiceless aspirated - voiced obstruents will probably stand out (Li 1973: 4). Another well-known example is the splitting of the Middle Chinese *Rù* 入 tone (the "entering tone" or "checked tone") into three or four subcategories in the Yue dialects, while in most of the Hakka and Min dialects the *Rù* tone only splits into two subcategories (Beijing Daxue 1995, 2003). Therefore, the splitting of the *Rù* tone usually helps distinguish a Yue dialect from a Hakka or a Min dialect (Norman 1988, 2003: 80). In these cases, examining the modern reflex of the Middle Chinese voiced obstruents and the evolution of the Middle Chinese *Rù* tone are useful diagnostic tests for dialect classification and comparison.

In short, with reference to the Middle Chinese sound system, fieldwork data of modern Chinese dialects have been recorded and sorted out according to the same criteria. The same foundation of dialect description in connection with Middle Chinese enables efficient comparisons among modern Chinese dialects.

2.2.3   Doublets and Multiple Layers of Phonological Forms

The traditional methodology also accommodates the study of the multiple layers of pronunciation in Chinese dialects. The literary and colloquial layers of pronunciation of the same morpheme forms doublets (and sometimes triplets) in Chinese dialects. Doublets, which have phonologically different forms but identical etymological origin, have a long history in China.

China has been a highly centralized nation for over two thousand years, ever since the unification of Qin (221 BC). As Li (2015: 592) points out, as a symbol of national unity, a standardized national language is "needed for effective governance, socioeconomic development, social advances through education, and the cultivation of shared cultural values." Since China's power centers (both political and cultural) were located in the North for the longest periods of time, the official, national language was naturally based on the Northern dialect in a majority of Chinese dynasties. According to Ho (2015: 149), this has been especially the case for about a millennium after the formation of Mandarin.[19]

It is also worth noting that, for much of China's history, Confucianism was the dominant and official ideology, roughly from the 2nd century B.C. through the early 20th century. As a result, during this long period of China's history, the study of ancient classics—especially Confucian ones—was particularly emphasized, along with all of the necessary reading, writing, and reciting skills. Without a strong competence in the literary language, it would be impossible for a Southern scholar to succeed in the imperial

---

[19] Mandarin was formed between the 12th and 14th centuries (Norman 1988: 48-49).

examination (科舉考試 *kējǔ kǎoshì*), which was often the only channel for social

mobility in ancient China as far as education is concerned.

To acquire abilities in both the Northern-dialect-based official language and in the

classic studies, Southerners had to rely on education. The linguistic consequence of this

formal education was the split of the literary and colloquial layers in the Chinese dialect,

particularly clear in the South. According to Ramsey's (1992: 38) definition, the

colloquial layer is "made up mostly of informal, everyday words," while the literary layer

contains "usually the more elevated terms of higher culture, which as a general rule came

into the dialect through the local tradition of reading literary texts." Some commonly seen

morphemes have differentiated pronunciations on the colloquial and on the literary

layers. In traditional terminology, this phenomenon is often referred to as 文白異讀

wénbái yìdú, "the differentiation of the literary and colloquial pronunciations."

Ramsey's definition, however, should not be misunderstood to mean that the

literary reading is rarely encountered in daily life. In fact, both ways of reading are

common in everyday speech. In fact, as Hashimoto (1973: 352) notes, the two layers are

"phonologically distinct in some systematic ways but are always associated with stylistic

differences." For instance, the morpheme *zhòng* 重 in Standard Cantonese has two

pronunciations. As a mono-morphemic adjective, *zhòng* 重 "heavy" reads [tʃʰɔŋ^{13}], which

is in the colloquial layer and reflects the native phonology of Cantonese; whereas in the

compound, bimorphemic adjective *zhòngyào* 重要 "important", it is unaspirated and

pronounced as [tʃɔŋ^{22}], which has a more formal, literary flavor.[20]

      While the colloquial layer presents the earlier forms of the spoken dialect, the

literary layer shows the later impact of literary education. These two pronunciations co-

exist in harmony in modern Southern dialects. The traditional methodology hence

reminds investigators of Chinese dialects to pay attention to the potential of dual or

multiple ways of pronouncing a single morpheme.

2.2.4   Drawbacks of the Traditional Methodology

      There are also some shortcomings associated with using the traditional approach.

First, the main focus is on the description and subsequent comparison of the literary layer

with regard to Middle Chinese, showing how dialect speakers *read* morphemes, but

critically, not necessarily how people *say* these words in daily life.

      Second, in the traditional description and comparison methodology it is also

possible that some of the morphemes collected are only rarely used in the local dialect.

This is especially the case for those dialects without a literary tradition. For instance,

when dialect speakers are asked to pronounce a morpheme that does not exist in the daily

vocabulary in the local dialect or that they do not know, they may make a guess. In either

scenario, the expectation to utter a pronunciation could push a speaker to make a

pronunciation out of reluctance, and the validity of such a pronunciation would be

---

[20] The tonal notation here follows Chao's (1930, 1980) system of tone numbers, which will be returned to with a more detailed introduction in §3.2.4. The tones "13" and "22" are the fifth and sixth tones in Cantonese, indicating low rising and low level contour, respectively.

questionable.[21] In addition, a complete registry of vernacular sounds may not be able to be elicited by reading morphemes alone. For example, I did not realize that there are syllables such as [piaŋ²²] "to hide" and [tʰiak⁴²] "to chase" in the Dapeng dialect, which were not covered by the standard morpheme list, until I heard native speakers say them casually in spontaneous speech.

In fact, some dialectologists have pointed out the drawbacks of relying solely on the Middle Chinese sound system and have begun to challenge the convention. Norman and Coblin (1995), for instance, have advocated a dynamic approach that is required for a "balanced and realistic study of Chinese dialect," which should be "a carefully balanced blending of the classical comparative method with the judicious use of written materials (for instance, rime books and rime tables)." They describe their approach as a combination of these philological sources with real linguistic data, which may be ignored if scholars only focus on individual morphemes in the written materials.

In particular, Norman and Coblin propose that a clearer classification of the Chinese dialects is the first step; then scholars can start to develop a better understanding of the major dialect groups. Next, one can conduct horizontal comparisons of modern Southern dialects, especially with those belonging to the same dialect group. This is a critical complement to the conventional, vertical comparisons of modern dialects with Middle Chinese. Both the clear classification of dialects and the close examination of

---

[21] This was not a serious problem in earlier dialect surveys. Back then it was not as difficult to find informants who received formal education in their local dialects. However, due to the successful promotion of Standard Chinese as the new medium of formal education in the past decades, it is more difficult to find informants who are literate in their local dialect, especially local dialects that do not serve as a regional *lingua franca*. Therefore, native speakers' shrinking ability in local pronunciation is a relatively new challenge for contemporary dialect surveys.

spoken language data provide researchers with a "safe ground" to employ the classical comparative methods.

This combined method is of particular importance in the investigation of the Southern dialects, where there exists a complex, multi-layered vocabulary (for instance, literary reading vs. colloquial reading), and where numerous large-scale waves of recurrent migration from the North took place throughout history. In addition to the diachronic examination of the literary language, as Norman and Coblin argue, only by carefully collecting and analyzing the complicated use of spoken language can dialectologists today reach a relatively realistic picture of a modern dialect and its relationship with others. The conventional methodology of dialect description, in light of their proposal, is insufficient due to the lack of spoken language data.

It should also be mentioned that Norman and Coblin's proposal has been gradually accepted, albeit in a critical manner, by the field of Chinese comparative dialectology. As Handel (2008) comments, in the past decades the centrality of the text-based reconstruction of Middle Chinese has been weakened in the field, while more emphasis is being put on the collection, analysis, and comparison of spoken-language data. While Norman and Coblin's new approach addresses the inadequacies of the traditional methodology, a fundamental limitation still seems to exist.

Apparently Norman and Coblin's framework is built based on a "clear classification of the Chinese dialects"; however, they do not articulate how exactly such a classification has been or will be achieved. As Ting (2003) has pointed out, their approach works well when comparing sub-varieties within the same dialect group but

does not seem to serve adequately the purpose of comparison across major dialects. If a classification scheme is the foundation of further, detailed comparison, as Norman and Coblin have argued, then a clear classification of major Chinese dialects grouping is critical in the first place, coming before comparisons within each dialect group.

2.3   Descriptive Methodology: Lexicon and Syntax

The traditional descriptions of the lexicon and syntax of Chinese dialects also have their strengths. As stated above, vocabulary lists are typically compiled according to semantic groups, and there is a similar set of simple sentences (e.g. interrogative, negative, double object, etc.) used to elicit their equivalents in the local dialect. These methods facilitate cross-dialectal comparisons. However, compared with the better-established methodology for describing phonology, lexical and syntactic descriptive methods have some shortcomings.

To be precise, it is generally the method of collecting data during fieldwork that is problematic, rather than its description in the dialect report. The essential problem lies in the material and method used to elicit answers from the informants. Both the vocabulary and sentence lists, in most cases, are written in characters (in Standard Written Chinese) prior to the fieldwork. Such lists are certainly helpful, but there is always an unintended danger if the dialect investigators ask questions like "how would you say *xxx* (a word) or *xxxxx* (a sentence) in the local dialect?" In even worse cases, the informant might be provided with the word or sentence list.

In either scenario, the informants could be influenced by the standard, written forms of the lexical or grammatical items. When this happens, the data collected may not reflect the local, colloquial words and structural patterns. Instead, what might be recorded is the informants' way of reading the characters in those prepared items.[22]

This problem can be illustrated by two examples, one from my own fieldwork in Dapeng and the other from the literature. During my fieldwork, in one of the sections where the informants were reading (aloud) a paragraph written in Standard Chinese, they had no problem pronouncing the morpheme *hé* 和 ("and") in the local dialect as [wɔ³¹]. However, based on my observation, they used another morpheme *tóng* 同 ("and") (pronounced as [tʰung³¹]) as the conjunction in their everyday speech. *Hé* 和 was rarely used as an alternative.

Another example, related to syntax, comes from Matthews' (1996) study of the phenomenon of ditaxia ("the co-existence of two syntactic alternatives, stratified by register and by social variables") in Cantonese. He points out (page 1275) that "some differences between Cantonese and Mandarin grammar are very subtle. Almost any Mandarin grammatical pattern can be used in Cantonese and be understood, but such locutions are often not idiomatic." For instance, "ŋɔ³⁵ pei³⁵ kʰøy³⁵ kuɔ³³" 我比佢高 (I-than-he-tall) and "ŋɔ³⁵ kou⁵³ kuɔ³³ kʰøy³⁵" 我高過佢 (I-tall-than-he) are both grammatical and have the same meaning, i.e. "I am taller than him." However, only the latter is regarded as colloquial, while the former, influenced by the Mandarin comparative

---

[22] The pronunciations of these morphemes are also important linguistic data, but the need to collect morpheme pronunciations has been fulfilled in the phonological investigation.

construction, is used "more by more educated and younger speakers, and in more formal registers" (Matthews 1996: 1278). In either case, given the influence of the standard and/or written Chinese, it seems less likely that the informants will spontaneously use the colloquial forms of either "and" or "I am taller than him."

This example shows the diverse, stratified nature of Chinese dialectal syntax. However, it used to be well accepted by early Chinese dialectologists that the grammar across Chinese dialects is essentially the same, a view espoused by Y. R. Chao's (1968: 13, as quoted below). As a result of this view, recording dialectal syntax tends to be neglected.

> "(Third,) it is in matters of grammar that the greatest degree of uniformity is found among all the dialects of the Chinese language. Apart from some minor divergence, such as indirect object before direct object in the Wu dialects and Cantonese –for which Mandarin (like English) has the opposite order, and slight differences in the order of the negative in potential complements in some of the southern dialects, and so on, and apart from differences in suffixes and particles for which, however, fairly close equivalents can be set up between dialects, one can say that there is practically one universal Chinese grammar."

In recent decades, linguists have argued against the so-called "universal Chinese grammar" (Yue-Hashimoto 1993, Matthews 1999, and Hashimoto 2008, among many others). Their research demonstrates that, contrary to Chao's claim, syntactic differences may exist in a great number of grammatical features across Chinese dialects (Yue-Hashimoto 1993), some of which are even of areal and typological significance (Hashimoto 2008). Since the 1980's, more attention has been drawn to the divergent

aspect of Chinese syntax (Kurpaska 2010), especially among non-Mandarin dialects, and now the academia has reached a general consensus on the heterogeneous nature of Chinese dialectal syntax.

A dialect report should address such diversity in Chinese syntax and record syntactic features of the dialect that show such diversity. A key question then, is which features should be recorded. Limited by volume, a dialect report normally does not include a comprehensive record of all syntactic features in the dialect;[23] rather, dialect investigators often include "distinctive" syntactic features only. According to Zhan et al. (1991) and You (1991), "distinctive" features refer to those that are different from the syntactic features in Standard Chinese (including both Putonghua, the spoken national language and the Standard Written Chinese). For instance, in Matthews' (1996: 1275) example, the word order of the comparative construction "ŋɔ^{35} kou^{53} kuɔ^{33} kʰøy^{35}" 我高過佢 (I-tall-than-he) in Guangzhou Cantonese should be regarded as a syntactic feature that does not exist in Standard Chinese.

Given the introductory nature of the dialect report, only focusing on this kind of distinctive feature is economical. More importantly, the primary purpose of the dialect report is to record and describe non-Standard Chinese dialects and to highlight their special characteristics—which precisely distinguish them both from the standard form of Chinese and from other dialects. For both reasons, lingering over syntactic features that are shared with Standard Chinese (and are hence "non-distinctive") is somehow

---

[23] Liu (2008) has proposed a comprehensive framework based on Comrie and Smith's (1977) typological questionnaire. This framework is extremely detailed and is therefore a great option for a monograph on dialectal syntax. For a dialect report, however, it is too voluminous.

redundant and may blur the focus of the dialect report. Overall, the dialect report maintains a good balance between the adequacy of revealing cross-dialectal diversity and the brevity of only providing distinctive syntactic features. This is another strength of the dialect report.

2.4   Fieldwork in the Dapeng Area

The previous sections discussed the strengths and weaknesses of the traditional methodology of Chinese dialect description. As stated above, the nature of dialect description is largely dependent on (if not entirely determined by) the methods of fieldwork. A well-established fieldwork method can outline a Chinese dialect with much clarity and accuracy, as in phonology; on the other hand, a problematic data-collection approach may lead to inaccuracies in the dialect description, as sometimes encountered in lexicon and syntax.

In the case of the undocumented Dapeng dialect, a revised fieldwork approach is needed for a more accurate description that aims to retain the strengths of the traditional fieldwork methods and overcome the weaknesses of the approach. In this section I introduce the refined fieldwork methods carried out in the Dapeng area, which combine traditional Chinese fieldwork and general sociolinguistic approaches.

The in-depth fieldwork was conducted from May to July, 2014, in the Dapeng area. With the assistance of the local museum and archive center, I gradually built up connections with middle-aged professionals in the Dapeng community. In the meantime, I visited the local senior activity center, where the senior Dapeng dialect speakers

gathered almost every workday afternoon to chat and dance. I met dozens of elders there and became friends with a few. Through a young social worker in the senior center I was also introduced to, and made acquaintance with, four young native speakers of the Dapeng dialect who were in their 20's.

Building relationships with people from different age groups was tremendously helpful in recording the local speech across generations and in observing the overall linguistic behavior of the Dapeng community. In addition to my network in the community, I also obtained some basic census data and other relevant demographic and historical information about the Dapeng area from the local museum, archive center, and government.

In the first part of the fieldwork design, "reading tasks", the approach was essentially a modification of the traditional methodology. I elicited and recorded the pronunciation of commonly used Chinese morphemes, which were listed according to sound categories in Middle Chinese. These morphemes, represented by Chinese characters, were drawn from two sources, including 1) the *Dialect Survey Character List* (方言調查字表 *Fāngyán diàochá zìbiǎo*, Zhongguo Shehui Kexueyuan Yuyan Yanjiusuo 1981), the standard word-list designed specifically for Chinese dialect surveys, and 2) a more recent character list in the *Linguistic Atlas of Chinese Dialects* (漢語方言 地圖集 *Hànyǔ fāngyán dìtú jí*, Cao 2008) and its handbook (漢語方言地圖集調查手冊 *Hànyǔ fāngyán dìtú jí diàochá shǒucè*, Beijing Yuyan Daxue 2003). From both sources I recorded approximately 2,700 commonly used morphemes altogether.

I also recorded the readings of a portion of the well-known ancient Chinese text for elementary literacy and moral education, *Three Character Classic* (三字經 *sānzì jīng*, written in the 13th century), as well as a part of the famous modern Chinese essay, *Tribute to the White Poplar* (白楊禮讚 *báiyáng lǐzàn*, written by Mao Dun 茅盾 in 1941). Both types of recording enabled dialect description that revealed phonological connections between the Dapeng dialect and Middle Chinese based on the historical phonological categories at the literary layer, that is, in formal reading.

With regard to lexicon and syntax, more adjustments to the traditional fieldwork methodology seemed necessary. Again, the method of data elicitation is particularly problematic in the process of lexical and grammatical data collection. Given these shortcomings, my revision of the traditional methodology mainly involved the elicitation of lexical items and grammatical patterns in the Dapeng dialect. The second and third parts of the fieldwork, "speaking tasks" and "spontaneous speech tasks", were designed to elicit lexical items and grammatical patterns from the Dapeng speakers at different consciousness levels. These tasks also complemented the reading tasks and contributed to the description of the Dapeng sound system at the colloquial layer, which reflects the local dialect in casual, everyday speech.

More specifically, in order to elicit colloquial words from my informants with more accuracy, I decided to limit their exposure to any standard or written forms of lexical items on my vocabulary list, which was based on the *Linguistic Atlas of Chinese Dialects* word list (Cao 2008). Instead of asking them to orally translate a spoken *lingua franca* item into a Dapeng item—which could potentially include undesirable features

from the *lingua franca*— I prepared a strategy of nonverbal elicitation for target items,

such as pictures, mainly for nouns, and rendered through performance other items such as

verbs and adjectives, asking the informants to then speak the words. For instance, I

imitated the sound of vomiting and elicited the colloquial word *ǒu* 嘔 (in Dapeng, [eu^{22}],

vomiting"), which appears as tù 吐 (a synonym of *ǒu*, the more frequently used form in

Standard Chinese) in the standard word list.

On occasions when informants were confused by my acting, I turned to explain

the *situation*—rather than the target word itself—in Cantonese, which most Dapeng

speakers perfectly understand. If the informant said a colloquial word, I would confirm

its being the target word by referring to its equivalent in Cantonese. For instance, to elicit

the verb "to chase (after)", I said "if a thief runs, then a cop will _____." The informant

immediately gave [t^{h}iak^{42}], which is a colloquial form. Then I asked whether it means

zeoi^{1} 追 in Cantonese, and received a positive answer. Most of the lexical items that I

recorded were successfully elicited by one of these three methods, and mainly by the first

two. Chapter 4 provides a complete list of those basic vocabulary items.

To collect grammatical patterns, I chose not to use a sentence list but to rely on

spontaneous speech only. In the "speaking tasks" part, I recorded story-telling using three

sources: 1) a picture book (*Frog, Where are You?* Mayer 2003), 2) a short silent film

(*The Pear Film*, Chafe 1980), and 3) a famous Aesop parable (*The North Wind and the*

Sun). I also recorded spontaneous narrations on several topics (Spring Festival

celebration, TV programs, the Dapeng history, etc.), as well as local people's casual

conversation.[24] Also in Chapter 4, grammatical patterns of the Dapeng dialect are concluded from these speaking tasks.

In addition to these reading and speaking tasks, I conducted a series of interviews, either formal or informal, with all participants at different stages of the fieldwork. I asked about Dapeng dialect speakers' biographical and linguistic background information, their language use in different social settings, as well as their attitudes towards the Dapeng dialect vis-à-vis the other three more widely spoken varieties of Chinese (i.e. the official language Putonghua, Cantonese, and Hakka).[25]

In total, I interviewed twenty native speakers of the Dapeng dialect, with the participants' age ranging from 22 to 84, roughly half male and half female. After all these interviews were finished, I recruited a young native speaker of the Dapeng dialect who was both literate and skillful in using a computer to transcribe all the audio recordings for future data analysis. Despite the presence of data from all ages, this dissertation in most parts only focuses on senior speakers of Dapeng over the age of 65 who speak a more conservative form of the local dialect, and leaves cross-generational investigations for future research.

---

[24] It should be pointed out that although the traditional methodology of describing a Chinese dialect also includes the collection of spontaneous speech, its purpose is more likely to record a piece of authentic speech than to provide data for formulating grammatical rules of the spoken language. In this regard, therefore, a spontaneous speech is not treated too differently from a lyric of a local ballad nor from a set of idioms in the local dialect.

[25] A complete list of interview questions is placed in Appendix A.

Chapter 3   Phonetics and Phonology

Following the convention of Chinese dialectology, this chapter describes the

phonetics and phonology of the Dapeng dialect in the format of the "dialect report." It

begins with a brief introduction of the informant's background and the research

procedure (§3.1). This chapter then describes the sound inventories of the Dapeng dialect

under the traditional framework, separating the syllables into initials, finals, and tones

(§3.2). §3.3 compiles a syllabary that puts together all legal Dapeng syllables, which are

formed as combinations of the 18 initials, 41 finals, and 5 tones as listed in §3.2. The

syllabary is complemented by the comprehensive list of homophonous morphemes in

Appendix C. This chapter also offers some general discussion of the Dapeng sound

system (§3.4).

3.1   Background: Informant and Procedures

The description in this chapter is based on the pronunciation of a senior native

speaker, Mr. L. He was 67 years old at the time of recording, and his pronunciation was

clearly audible. He was born and raised in a Dapeng-speaking family and later married

another native Dapeng speaker, raising their children solely in the local dialect. Before

retirement, Mr. L had worked as a taxi driver in the local community and had never lived

outside of Dapeng for a long period of time.

Mr. L finished middle school and was therefore literate. Teachers back when he was a student came from local and adjacent areas and could be Dapeng, Cantonese, or Hakka speakers. He remembered in his school having exposure to all these varieties, plus some Putonghua. Mr. L and his wife watched TV at night, watching programs in both Cantonese and Putonghua. (There were no TV stations broadcasting programs in either Hakka or the Dapeng dialect in their area, at least at the time of recording.)

Mr. L used the Dapeng dialect in most daily settings: both at home (to his spouse, children, and grandchildren) and in the neighborhood (to other senior residents). According to both a self-evaluation and his peer's assessment, Mr. L spoke an authentic Dapeng dialect. In addition to the local speech, he could also speak fluent Cantonese and Hakka, and was also capable (albeit less fluent) in Putonghua. He only switched to the other varieties of Chinese occasionally in order to accommodate his non-Dapeng interlocutors. For instance, he spoke Putonghua to his tenants, migrant workers who came from Mandarin speaking areas

The majority of the recordings with Mr. L were conducted in a quiet room, with only him and the interviewer present. These sessions included reading tasks, speaking tasks, and some spontaneous speech. They were all recorded by a Tascam DR-07 MK II recorder at 44,100 Hz, in the 24-bit WAV format. A lavalier microphone was used to reduce the distraction caused by the audio recorder. A small portion of the recordings, mostly daily conversation with other local people, were conducted in public places with the consent of all participating speakers. This data was recorded with an Etekcity digital audio pen recorder in the 128K bps WAV format.

3.2   Sound Inventories

The description of the Dapeng sound inventories follows the traditional initial-
final dichotomy. As introduced in §2.1, the "initial" is the initial consonant of a syllable;
the "final" refers to the remaining segments, including the medial vowel, the nucleus, and
the ending, among which only the nucleus is compulsory. Tones are also introduced. In
this section, initials, finals, and tones in the Dapeng dialect are put into tables. Necessary
notes are also included. Unless specified otherwise, this section follows the "dialect
report" and uses broad phonetic transcription.

3.2.1   Syllable Structure

The Dapeng dialect shares the same syllable structure with most other Chinese
dialects, viz. CGVX. In this structure, C is either a consonant or a glide, G is a glide, V is
a vowel or a syllabic consonant, and X is a nasal/stop consonant or a glide. C corresponds
to "initial" and GVX corresponds to "final." A syllable must have a V, while other
segments are optional (Yue-Hashimoto 1972: 87-88).

Based on the general syllabic structure of Chinese as introduced in §2.2.1,
possible candidates in the Dapeng dialect at each position are listed below each
component. It should also be pointed out that not all possible combinations of C-G-V-X
are legal syllables in the Dapeng dialect. The restrictions will be shown in detail in §3.3
and in Appendix C, which lists all combinations of initials, finals, and tones in the
Dapeng dialect, based on the fieldwork collected by this author.

Tone (42, 31, 55, 22, 54)			
Initial	**Final**		
	Medial	Nucleus[26]	Ending
Consonant/glide	*Glide*	*Vowel*	*Glide or nasal/ stop consonant*
p, pʰ, f, m, t, tʰ, n, l, ts, tsʰ, s, k, kʰ, ŋ, h, i̯, u̯	i̯, u̯	a, ɐ, ɛ, ɔ, i, u	i̯, u̯, p, t, k, m, n, ŋ

Syllable Structure of the Dapeng Dialect

Before the description continues, it is crucial to point out that glides in the Chinese language can be predicted within the structure of a syllable. For instance, the phoneme /i/ could be used for the syllabic segment as well as its non-syllabic counterpart depending on whether it occupies the nucleus position of the syllable: it is the vowel [i] in the nucleus position and the non-syllabic [i̯] in other positions. Since whether a segment is [+syllabic] or [-syllabic] in a syllable is usually predictable, the diacritic for [-syllabic] in [i̯] and [u̯] can be omitted, and diphthongs such as [ai̯] and [au̯] can simply be transcribed phonemically as /ai/ and /au/.

The avoidance of diacritic marks is also in accordance with Chao's (1934: 390) general suggestion on the choice of symbols in phonemic transcription. This suggestion has been taken in the majority of the recent descriptive works in Chinese dialectology, such as Beijing Daxue (1995, 2003), Yuan (2001), and Zhan (2002). For both reasons of

---

[26] The syllabic nasal [m̩] can also occur as a nucleus, but it has to occupy the entire syllable by itself. Since [m̩] is not in combination with any other segments, it is not put in the table. Also, [ɐ] are [ɛ] are allophones of the phoneme /ɐ/. Therefore, there are only five phonemes in the Dapeng dialect. This will be discussed in more detail in §3.2.3.

predictability and following the convention, in the subsequent sections and throughout the rest of this dissertation, the Dapeng glides in both medial and ending positions will be broadly transcribed as /i/ and /u/ without using diacritic marks. In initial position, they are transcribed as /j/ and /w/, respectively.

3.2.2 Initials

There are 18 initials in the Dapeng dialect, as tabulated in Table 1. This table also contains all consonantal phonemes in this dialect. While all these phonemes can appear in syllable-initial position according to the definition of "initial", only six of them can possibly occur in syllable-final position as a consonantal ending: /-p/, /-t/, /-k/, /-m/, /-n/, and /-ŋ/, as is revealed in the table of finals in §3.2.1. The only initial in Table 1 that is not a consonant or glide is "○-", the "zero initial" (零聲母 *língshēngmŭ*).

	Plosive		Fricative	Nasal	Lateral	Approximant	-
	Unaspirated	Aspirated					
Labial	p	pʰ	f	m		w	
Dental	t	tʰ		n	l		
Alveolar	ts	tsʰ	s				
Palatal						j	
Velar	k	kʰ		ŋ			
Glottal			h				
-							○

Table 1. The Initials in the Dapeng Dialect

41

The table of initials distinguishes places of articulation in rows and manners of articulation in columns. Some notes should be made, each addressing some nuances that are not revealed in the broad phonetic transcription in Table 1.

Since the bilabial initials (/p-/, /pʰ-/, /m-/, and /w-/) and the labiodental initial /f-/ are in complimentary distribution, in Table 1 the two groups are combined as "labial initials"; /w-/ can be alternatively placed in the velar row due to the labial-velar coarticulation. The initials /n-/ and /l-/ from the dental group maintain their distinction in the Dapeng elders' speech; however, the younger generation (below 40's) seems to have started a merger from /n-/ to /l-/. Based on Mr. L's pronunciation, Table 1 presents a more conservative initial system, which clearly distinguishes /n-/ from /l-/.

A syllable with the zero initial starts smoothly with the nucleus. Figures 1 and 2 are two examples: The spectrograms of the syllable [un^{42}] (安, "safe") and the syllable [in^{31}] (言, "speech") show no clear obstruction in the vocal tract at the beginning of the pronunciation.[27] However, despite the lack of acoustic signal, the current study follows the tradition of Chinese dialect description and assigns the zero initial "○-" to the syllable-initial position, so that syllables without an audible consonant onset can still be analyzed under the same initial-final framework.

---

[27] And therefore, the onset is less likely a glottal stop initial.
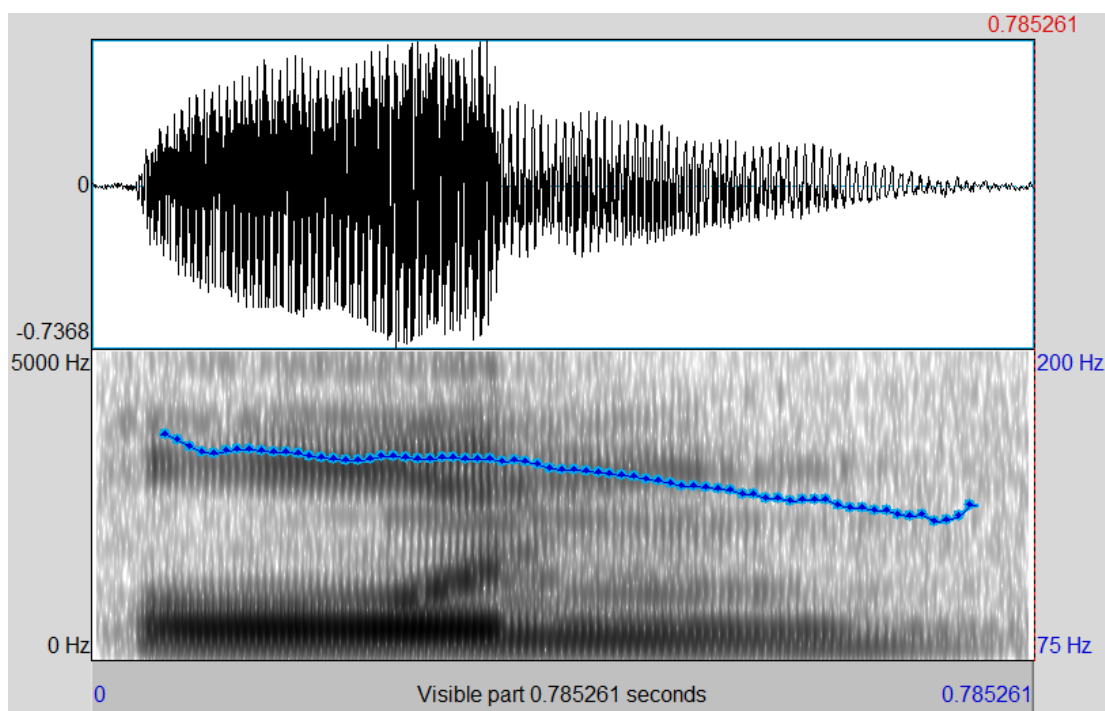
Figure 1. Spectrogram of [un⁴²] (安, "safe")
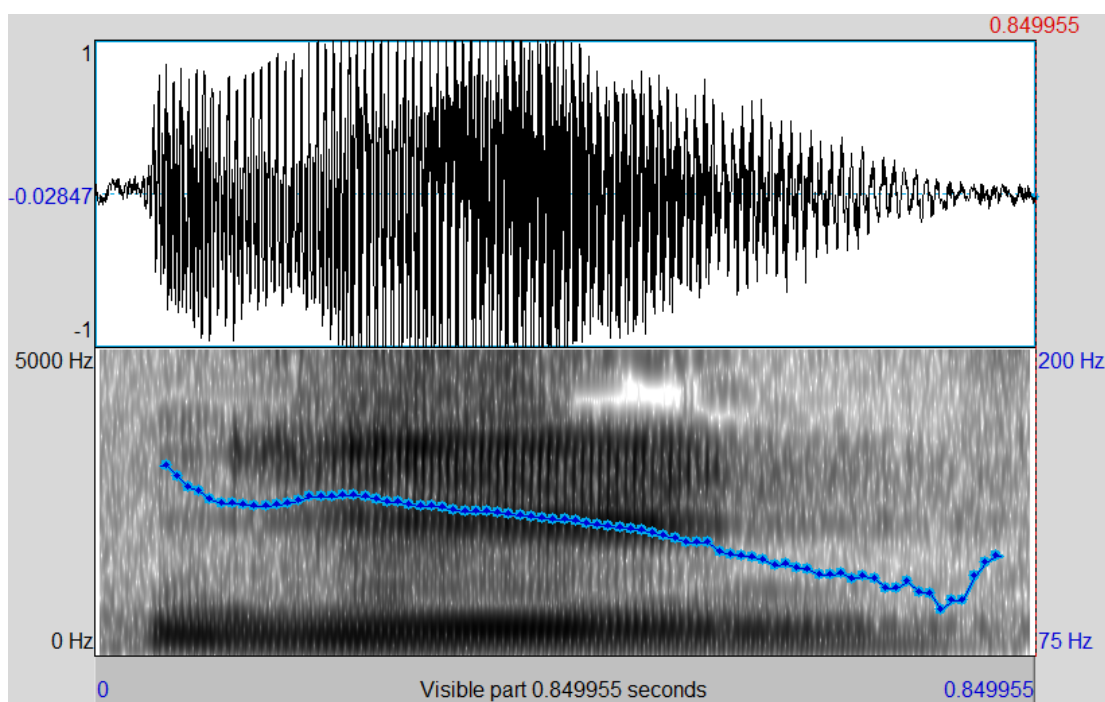
Figure 2. Spectrogram of [in³¹] (言, "speech")

43

3.2.3   Finals

The Dapeng dialect has a relatively simple vowel inventory. Figure 3 shows that there are only six vowels in this dialect. All six vowels can occur as nuclei, while only /i/ can occur in the medial or ending position, both optional in the Dapeng syllable structure. The other high vowel, /u/, can only occur as an ending.[28] Besides high vowels, the Dapeng ending can also be nasal or stop consonants. Unlike at the initial position, all stop consonants at the ending position are unreleased.

Table 2 presents all of the 41 finals in the Dapeng dialect, each being a combination of vowels and glides or consonant at the three positions: medial, nucleus, and ending. They are tabulated in columns according to the nucleus. In each row, the main vowel is in a type of combination with the medial and/or ending.
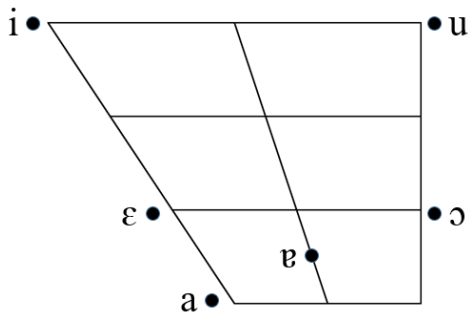
Figure 3. The Dapeng Vowel Inventory

---

[28] Again, the narrow, phonetic transcriptions for /i/ and /u/ in the medial or ending position are [i̠] and [u̠], respectively.

V (nucleus) alone	a		ɔ	i	u	-
V + ending [-i]	ai	ɐi			ui	
V + ending [-u]	au	ɐu		iu		
V + ending [-m]	am	ɐm		im		
V + ending [-n]	an	ɐn		in	un	
V + ending [-ŋ]	aŋ	ɐŋ	ɔŋ		uŋ	
V + ending [-p]	ap	ɐp		ip		
V + ending [-t]	at	ɐt		it	ut	
V + ending [-k]	ak	ɐk	ɔk	ik	uk	
medial [-i-] + V	ia	iɛ	iɔ			
medial [-i-] + V + ending [-ŋ]	iaŋ		iɔŋ			
medial [-i-] + V + ending [-k]	iak		iɔk			
syllabic consonant						m̩

Table 2. The Finals in the Dapeng Dialect

Among all these finals, /-iu/ and /-ui/ are the two that might cause confusion.
Given [i] also as a medial in Dapeng and [u] as another medial commonly seen across
Chinese dialects, it seems possible in theory to treat /-iu/ as "medial [i̯] + nucleus [u]" and
/-ui/ as "medial [u̯] + nucleus [i]." However, the current study treats these two finals as
"nucleus + ending," that is, /-ui/ as "nucleus [u] + ending [i̯]" and /-iu/ as "nucleus [i] +
ending [u̯]." From a phonetic perspective, the first vowel in both finals sounds heavier
and longer, and the second vowel is in comparison lighter and shorter. Meanwhile, the
first vowel is the tone-bearing unit in both combinations. Therefore, the first vowel
should be treated as the main vowel in both /ui/ and /-iu/.

Another observation one can make from Table 2 is that neither [ɐ] nor [ɛ] can
occur in the nucleus position by itself; as nuclei they have to take other segments as
medials and/or endings. That is, [ɐ] has to precede an ending, either a glide or a
nasal/stop consonant, while [ɛ] can only co-occur with the medial [i̯]. Since their

45

occurrences are in complementary distribution with each other, [ɐ] nor [ɛ] can be treated

as allophones of the phoneme /ɐ/. According to Table 2, the phoneme /ɐ/ is realized

phonetically as [ɛ] when it occurs after the medial [i̥], and as [ɐ] in all other instances,

that is:

<div style="text-align: center;">

			Example
/ɐ/	→	[ɛ] / i̥ _	tsiɛ^{35} 姐 "older sister"
	→	[ɐ] / elsewhere	hɐŋ^{42} 兄 "older brother"

</div>

It should also be noted that the Dapeng dialect, like many Cantonese dialects,

contains the syllabic nasal [m̩]. That is, the nasal consonant [m] can constitute a syllable

without combination with any other segments. For instance, the pronunciation of the

morphemes *wǔ* 五 ("five") and *wǔ* 唔 (the colloquial negative marker) in Dapeng are

both [m̩], bearing a high rising tone and a low falling tone, respectively.

3.2.4   Tones

The Dapeng dialect has five tones. There are three falling tones: Tone 2 (low

falling), Tone 1 (mid-falling), and Tone 5 (high falling). Of the remaining two, one is

high rising (Tone 3) and the other is low level (Tone 4). All the tones are listed in Table 3

with some example morphemes, and their contours are visualized in Figure 4.

The tonal notation here and across the entire dissertation follows Chao's (1930,

1980) system of tone numbers, which has long been adopted in Chinese linguistics. In

Chao's system, the tone height is indicated by numbers "1" through "5", with "1" for

lowest pitch and "5" for highest pitch. Tone contours can be transcribed by a sequence of

these numbers, marking starting and ending points (and turning points also for bidirectional contour tones, such as a dipping tone). Take the Standard Mandarin Chinese tonal system as an example: Tone 1 is high level, transcribed as "55"; Tone 2 is high rising, "35"; Tone 3 is mid-dipping, "214"; and Tone 4 is high falling, "51".

Tone	Description	Tone numbers	Examples
Tone 1	Mid-falling	42	詩 "poem" [si^{42}], 呼 "shout" [fu^{42}], 刷 "brush" [ts^{h}at^{42}]
Tone 2	Low falling	31	時 "time" [si^{31}], 湖 "lake" [fu^{31}]
Tone 3	High rising	35	死 "die" [si^{35}], 虎 "tiger" [fu^{35}]
Tone 4	Low level	22	四 "four" [si^{22}], 富 "rich" [fu^{22}]
Tone 5	High falling	54	樹 "tree" [si^{54}], 父 "father" [fu^{54}], 疾 "sickness" [ts^{h}at^{54}]

Table 3. The Tones in the Dapeng Dialect
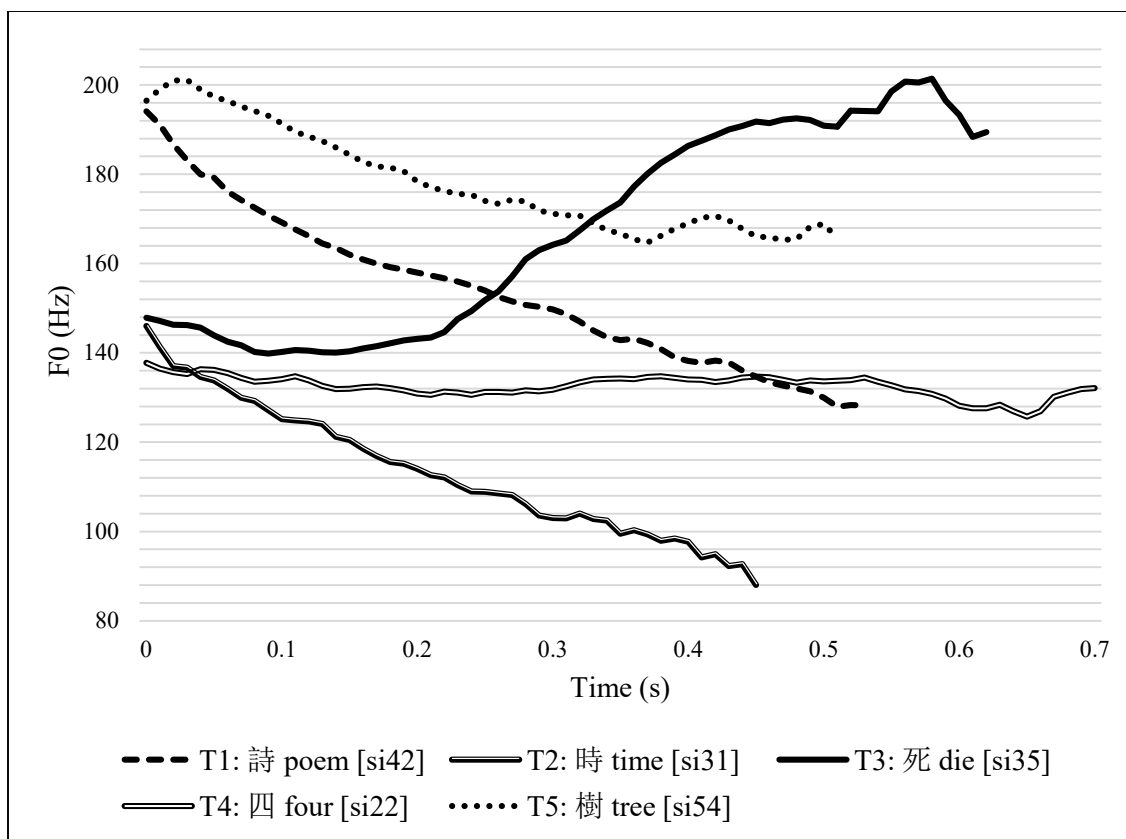
Figure 4. Contours of Dapeng Tones

The measurement of the Dapeng tones was based on 90 commonly used morphemes, with 20 morphemes from the Middle Chinese *Yángshǎng* 陽上 tonal category[29] and 10 each from other seven tonal categories. Mr. L's pronunciation of these morphemes were analyzed by the Praat software (Boersma and Weenink 2013). All morphemes was pronounced in isolation.

---

[29] The reason for treating the *Yángshǎng* category somewhat differently is that a pilot study showed a more irregular correspondence between that Middle Chinese tonal category and the modern Dapeng tones. Due to such complication, the *Yángshǎng* category had twice as many morphemes included in the current analysis than all the other tonal categories.

For each of the 90 tokens, the f_0 (fundamental frequency) was measured at five locations of the sonorous portion of each syllable: the onset (beginning, 0%), first-quarter point (25%), midpoint (50%), third-quarter point (75%), and the offset (end, 100%). Within each Middle Chinese tonal category except *Yángshǎng*, the average f_0 at each measure point was calculated across all 10 tokens; the *Yángshǎng* categories were first sub-categorized into three subgroups of tones by ear: high rising, high falling, and low level. Then the average f_0 of each subcategory was calculated.

The tone contour of each tonal category was then produced with a line that connects all five points. Based on the observation of the overall trends of these lines, all the tone contours were grouped into five modern tonal categories: high falling, mid-falling, low falling, high rising, and low level. That is, some Middle Chinese tonal categories have merged with others.[30]

In order to adjust the heights of the modern Dapeng tonal categories into Chao's (1930, 1980) five-scale system, the maximum and minimum f_0 values among all the averaged f_0 values were found across all tonal categories. The range between the maximum and minimum f_0 values was then divided into 5 equal levels, with the minimum being 1 and the maximum being 5. All the other values falling in between were adjusted according to the ratio. By converting the averaged f_0 values at each measure point into tone values in the 1-5 range, the approximate tone values for each Dapeng tonal categories were proposed, as shown in Table 3 and Figure 4 above.

---

[30] The topic of the correspondence between the Middle Chinese and modern Dapeng tonal categories will be presented in more detail in a later section in this chapter.

It should also be mentioned that the tones in the Dapeng dialect do not undergo tone sandhi.[31] One of the subjects (male, 59 years old) participated in an *ad hoc* investigation during the fieldwork, which was dedicated specifically to the study of tone sandhi in Chinese dialects. The subject was instructed to read through a list of disyllabic words from the *Handbook for the Linguistic Atlas of Chinese Dialects* (Beijing Yuyan Daxue 2003: 34). This list includes in total 196 disyllabic words, which include all possible tone combinations in modern Chinese dialects. Out of the 196, 191 of them (more than 97%) do not show any tone sandhi, i.e. both syllables in all these words are read with the same tones as their citation forms, while the remaining 5 cases of tone change randomly occur in different tone combinations. That is to say, tone sandhi is very rare and exceptional in the Dapeng dialect. There are occasionally isolated cases of tone sandhi, but overall there is no clear pattern.

3.3   Syllabary

Actual Dapeng syllables are formed as combinations of the previously listed 18 initials, 41 finals, and 5 tones. The following syllabary in Table 4 shows all possible occurrences, marked with "+," of syllables in the Dapeng dialect. If certain syllables are only observed in the colloquial language (which means no original morphemes are detectible—see §4.2 for more details), they are marked with "*". Illegal syllables are

---

[31] The most well-known tone sandhi in the Chinese language is probably the rule in Putonghua that when a 3rd tone (dipping) is followed by another 3rd tone, the first tone is pronounced with the 2nd tone (rising). For instance, nǐhǎo [3rd+3rd] becomes níhǎo]2nd+2nd] (你好 "hello").

unmarked. Initials are arranged in the first row, with bold lines separating different places of articulation. Finals and tones are listed in the first two columns.

		p	pʰ	f	m	w	t	tʰ	n	l	ts	tsʰ	s	j	k	kʰ	ŋ	h	○
a	42	+		+		+		+			+	+	+		+	+		+	+
	31		+		+	+			+		+	+	+				+	+	
	35	+			+		+		+				+		+		+		
	22	+	+	+	+						+		+		+			+	
	54		+		+	+						+	+				+	+	
ai	42					+					+	+	*		+				+
	31		+		+	+							+	+			+	+	
	35	+							+				+	*					+
	22	+	+	+	+		+	+	+		+	+	+		+				
	54		+		+	+		+	+	+		+					+	+	
au	42	+		+			+	+			+	+	+		+				
	31		+	+	+			+		+		+					+	+	
	35	+	+	+			+	+	+	+	+	+			+	+		+	+
	22	+	+	+			+	+			+	+	+		+	+	+	+	+
	54	+	+	+	+			+	+			+	+				+	+	
am	42						+	+				+	+		+		*	+	
	31							+	+	+		+					+	+	
	35						+			+	+	+			+	+			
	22						+	+		*					+			+	+
	54									+		+						+	
an	42	+	+			+	+	+				+	+		+				
	31				+	+		+	+	+		+					+	+	
	35	+				+		+			+	+	+		+				
	22				+		+	+			+	+	+		+		+		
	54		+		+			+	+	+		+					+	+	
aŋ	42	+	*								+	+	+		+			+	
	31		+		+	+					+	+							
	35									+			+			+			
	22		+		+								+		+				
	54			+								+					+		
ap	42						+	+				+			+				+
	31																		
	35																		
	22																		
	54									+		+	+		+			+	

Continued

Table 4. The Dapeng Syllabary

Table 4 continued

		p	pʰ	f	m	w	t	tʰ	n	l	ts	tsʰ	s	j	k	kʰ	ŋ	h	○
at	42	+		+		+					+	+	+		+				+
	31																		
	35																		
	22																		
	54		*	+	+	+		+		+	+	+							
ak	42	+	+								+	+			+			+	
	31																		
	35																		
	22																		
	54	+			+	+		+		*		+	+				+		
ia	42												+						
	31											+							+
	35												+						+
	22										+		+						+
	54				*							+							+
iaŋ	42						+	+				+	+		+			+	
	31		+		+							+							+
	35	+						+			+	+	+		+				+
	22	+								+			+		+				
	54		+		+		*			*		+				+			
iak	42	+	+					+					+						+
	31																		
	35																		
	22																		
	54							+								+			
ɐi	42		+	+		+	+	+				+	+		+	+			
	31					+		+	+	+		+				+	+	+	
	35				+	+	+	+		+			+		+	+			
	22			+		+	+	+			+	+	+		+	+			
	54		+			+		+		+						+	+	+	
ɐu	42	*			*			+			+	+	+	+	+			+	+
	31			+	+			+		+		+	+	+		+	+	+	
	35		+	+	+		+		+		+	+	+	+	+		+	+	
	22						+			+	+	+	+	+	+	+	+	+	+
	54			+	+			+		+		+		+		+		+	

Continued

53

Table 4 continued

		p	pʰ	f	m	w	t	tʰ	n	l	ts	tsʰ	s	j	k	kʰ	ŋ	h	○
ɐm	42										+	+	+	+	+	+			
	31									+		+		+		+		+	
	35		+				*				+		+	+	+				
	22										+		+		+				
	54		+										+	+					
ɐn	42	+		+	+	+	+	+			+	+	+	+	+	+	+	+	
	31			+	+	+		*		+		+	+	+		+	+	+	
	35		+	+			+			*	+	+	+	+	+	+		+	
	22		*	+							+		+	+		+			
	54		+	+	+	+				+		+	+	+	+	+		+	
ɐŋ	42	+	+				+				+	+	+	+	+	+		+	
	31				+			+	+	+		+	+	+		+		+	
	35	+	+				+	+	*		+				+	+	*		
	22	+	+			+	+				+	+	+		+			+	
	54				+			+		+			+	+				+	
ɐp	42									+		+	+		+	+		+	
	31																		
	35																		
	22																		
	54									+		+	+	+		+			
ɐt	42	+	+	+							+	+	+	+	+	+	+		
	31																		
	35																		
	22																		
	54			+	+					+		+	+		*	+		+	
ɐk	42	+	+				+				+		+	+		+	+	+	
	31																		
	35																		
	22																		
	54				+					+		+	+	+					
iɛ	42																		
	31															*			
	35										+	+							
	22										+								
	54																		

54

Table 4 continued

		p	pʰ	f	m	w	t	tʰ	n	l	ts	tsʰ	s	j	k	kʰ	ŋ	h	○
ɔ	42	+			+	+	+	+				+	+		+	+			
	31		+		+	+		+		+			+				+	+	
	35			+	+		+	+			+	+	+		+			+	
	22	+		+									+		+	+	+		
	54				+	+		+	+				+				+	+	
ɔŋ	42	+		+		+	+	+			+	+	+		+	+		+	
	31		+	+	+	+		+	+	+		+	+			+		+	
	35	+		+	+	+	+	+			+	+	+		+	+			
	22			+			+				+	+	+		+	+	+	+	
	54				+	+				+	+	+			+			+	
ɔk	42	+			+	+		+		+	+	+			+	+		+	+
	31																		
	35																		
	22		+																
	54		+		+	+				+		+					+	+	
iɔ	42							*				+						+	
	31									+						+			
	35						+								+				
	22																		
	54									+									
iɔŋ	42										+	+	+		+	+		+	+
	31								+	+		+				+			+
	35									+	+		+			+		+	+
	22										+		+					+	+
	54									+	+	+							+
iɔk	42						+				+				+	+			+
	31																		
	35																		
	22																		
	54																		+
i	42	+		+			*				+	+	+		+	+		+	+
	31		+	+	+				+	+		+	+			+			+
	35	+	+	+					+	+	+	+	+		+	+	+	+	+
	22	+	+		+				+	+	+	+	+		+	+		+	+
	54		+		+			+	+	+		+	+			+	+		+

Continued

55

Table 4 continued

		p	pʰ	f	m	w	t	tʰ	n	l	ts	tsʰ	s	j	k	kʰ	ŋ	h	○
im	42							+	+		+	+				+		+	+
	31							+	+	+			+		+	+			+
	35						+						+		+			+	+
	22						+		+	+	+				+			+	+
	54						*		+	*		+							+
in	42	+	+				+	+	*		+	+	+		+			+	+
	31				+			+	+	+		+	+			+		+	+
	35				+		+			+	+	+	+		+			+	+
	22	+	+		+		+	+	+		+	+	+		+			+	+
	54		+		+			+	+	+		+	+			+		+	+
ip	42						+	+			+	+			+			+	
	31																		
	35																		
	22																		
	54							+	+	+		+				+		+	+
it	42	+						+			+	+	+		+	+		+	+
	31																		
	35																		
	22																		
	54				+		+	+		+	+	+	+			+	+	+	+
ik	42	+					+												
	31																		
	35																		
	22																		
	54												+						
iu	42	+	+				+	+			+	+	+		+				+
	31		+		+			+		+		+	+			+			+
	35	+			+					+			+		+				
	22		+				+	+			+		+		+				+
	54				+		+	+	+	+		+				+			+
u	42			+			+				+	+	+		+				+
	31		+	+	+			+	+	+									
	35	+	+	+	+		+	+	+	+	+		+		+				
	22	+	+	+	+			+			+	+	+		+				
	54		+	+	+			+	+	+									+

Table 4 continued

		p	pʰ	f	m	w	t	tʰ	n	l	ts	tsʰ	s	j	k	kʰ	ŋ	h	○
ui	42	+	+	+			+	+			+	+	+			+		+	+
	31		+		+			+		+		+	+				+	+	
	35			+				+		+	+	+	+		+	+			
	22	+	+		+		+	+		+	+	+	+		+				
	54			+	+			+	+	+		+	+	+			+	+	+
un	42		+	+											+			+	+
	31		+	+	+					+								+	
	35														+			+	+
	22		+	+	+										+	+			+
	54		+	+	+													+	+
uŋ	42			+			+	+			+		+	+	+			+	
	31		+					+	+			+		+		+		+	
	35						+	+			+			+					
	22						+	+			+		+		+				
	54			+				+	+				+	+		+			
ut	42		+	+											+			+	
	31																		
	35																		
	22																		
	54	+	+		+			+										+	+
uk	42		+	+			+			+	+	+	+		+	+		+	
	31																		
	35																		
	22																		
	54			+	+			+		+		+		+		+			
m̩	42																		
	31																		+
	35																		+
	22																		+
	54																		+

57

As a complement to the syllabary, a typical dialect report also sorts out all Dapeng morphemes according to their phonetic form. Homophonous morphemes must have the initial, the final, and the tone all identical. Morphemes, whenever possible, will be represented by characters. A comprehensive list of homophonous morphemes is in Appendix C.

3.4   Diachronic and Synchronic Comparison

This section discusses some of the distinctive features in the Dapeng sound system, in addition to the previous brief introduction to sound inventories. These features will be addressed from both diachronic and synchronic perspectives. As introduced in §2.2.2, the diachronic comparison is usually between the modern dialect in question and the reconstructed Middle Chinese sound system, which shows an overall regular relationship to the majority of modern Chinese dialects (Norman and Coblin 1995: 582). According to the initial, final, and tonal categories in Middle Chinese, modern dialects are analyzed, and the historical development of the modern dialects can be suggested based on patterns of sound correspondence between the ancient categories and their modern reflexes.

As stated in §2.2, the diachronic comparison also enables synchronic comparison across modern Chinese dialects. That is, with reference to the Middle Chinese sound system, the vast majority of modern Chinese dialects have been described and analyzed under the same framework, and most of the dialectological fieldwork data have been sorted out according to the same criteria. In this section, the Dapeng dialect is also

compared with three of the major modern Chinese dialects based on the initial, final, and tonal categories in the Middle Chinese sound system.

Specifically, the comparison will show differences either between the Dapeng dialect and Putonghua or between Dapeng and Cantonese and/or Hakka will be addressed. However, the two types of differences do not have the same weight in the discussion to follow due to different predictability. As a Cantonese-Hakka mixed dialect, it is not surprising that the Dapeng dialect shares many non-Mandarin features with other Southern Chinese dialects, Cantonese and Hakka included. Those predictable features, therefore, will only be mentioned briefly in this section. In contrast, features that show difference from either or both of the major source dialects will be highlighted. Examples will be provided whenever necessary to demonstrate the second type of difference.

In this section and beyond, all Cantonese and Hakka materials are drawn from the following sources: Yue-Hashimoto (1972), Hashimoto (1973), Gao (1980), Li and Zhang (1992), Chen (1993), Beijing Daxue (1995, 2003), Zhan (2002), and National Taiwan University et al. (the Xiaoxuetang 小學堂 database). Unless specified otherwise, by "Cantonese" and "Hakka" I refer to their representative dialect, Guangzhou or Hong Kong Cantonese and Meixian Hakka, respectively.

It should also be noted that using the representative Cantonese and Hakka dialects does not suggest that the original contributing dialects of the early Dapeng were actually these two dialects. Given the lack of the earlier, original contributing dialect data, it is more practical in the current study to compare the modern Dapeng dialect with contemporary Cantonese and Hakka, represented by Guangzhou or Hong Kong

Cantonese and Meixian Hakka, respectively. The justification of this comparison is further discussed in §5.5.

3.4.1   Initials

Similar to the majority of the Cantonese and Hakka dialects, the Dapeng dialect only has one set of sibilants, namely, alveolar sibilants: /ts-/, /tsʰ-/, and /s-/. These initials come from the Middle Chinese Jīng 精, Zhī 知, Zhuāng 莊, and Zhāng 章 initial groups.[32] In Putonghua as well as many other Northern dialects, in contrast, the modern reflexes of these initial groups not only include one set of aveolar sibilants but also include one set of retroflex sibilants: /tʂ-/, /tʂʰ-/, and /ʂ-/.

The Dapeng dialect also preserves a clear distinction between the Middle Chinese Jīng 精 and Jiàn 見/Xiǎo 曉 initial groups. Today the Jīng 精 initial group corresponds to the Dapeng alveolar initials (/ts-/, /tsʰ-/, and /s-/), and the Jiàn 見/Xiǎo 曉 initial groups corresponds to velar initials (/k-/, /kʰ-/, and /h-/). This feature is shared by most of the Cantonese and Hakka dialects; in Putonghua, by contrast, such a distinction has been

---

[32] Middle Chinese initials can be put into groups based on the places of articulation. Those initial groups are called *Shēngzǔ* 聲組. For instance, the five initials Jīng 精 *ts-, Qīng 清 *tsʰ-, Cóng 從 *dz-, Xīn 心 *s-, and Xié 邪 *z- can combine as the Jīng 精 initial group according the shared place of articulation, alveolar. The Jīng 精 initial group (Jīngzǔ zì 精組字) is also called 齒頭音 *Chǐtóu yīn* in the traditional terminology, which means "alveolar." This dissertation follows Wang Li's reconstruction system of Middle Chinese (Wang 1980: 50-51, 1987: 281). His reconstruction is based on Zhu Ao's *fanqie* 朱翱反切, which reflects the real speech in the late Tang Dynatsty. *Fanqie*, literally meaning "turning and cutting", is an initial-final dichotomy used as a philological tradition to record the pronunciation of Chinese characters since the late Han Dynasty (Chan 1980: 18). For more details concerning *fanqie*, see Norman's (1988: 27-28) introduction.

basically lost, with both Jīng 精 and Jiàn 見/Xiǎo 曉 initial groups pronounced alveolar in a good portion of morphemes from these groups.

Another notable feature of the Dapeng initials is the velar nasal /ŋ-/. The main source of the Dapeng /ŋ-/ is the Middle Chinese Yí 疑 initial *ŋ-, with the Rì 日 initial *r- is the secondary source. This is in overall correspondence to both Cantonese and Hakka. In Putonghua, the Yí 疑 initial has developed to either zero /○-/ or glide initials (/j-/, /w-/),[33] and the velar nasal initial *ŋ- is no longer preserved.

The most striking feature of the Dapeng initial system is the development pattern of Middle Chinese voiced obstruents. The reconstruction of Middle Chinese indicates a three-way distinction in manner in the initial obstruents: It has voiceless unaspirated, voiceless aspirated, and voiced obstruents. This three-way contrast still exists in some modern varieties of Chinese, for instance, the Wu dialect (Li and Xiang 2009: 182). In contrast, the voiced obstruents have been lost in many others, including Putonghua, Cantonese, and Hakka. They have merged with either voiceless unaspirated or voiceless aspirated obstruents in these varieties of Chinese.

The development pattern of the voiced initials is usually influenced by tonal categories. For instance, in both Putonghua and Cantonese they became aspirated in syllables that have the Middle Chinese *Píng* 平 ("level") tone and unaspirated in the other tones (*Shǎng* 上, *Qù* 去, and *Rù* 入, altogether called the *Zè* 仄 ["oblique"] tones). In other varieties of Chinese such as Hakka and Gan, by contrast, the set of voiced initials

---

[33] Alternatively, the glide initials /j-/ and /w-/ can also be treated as the medials /-i-/ and /-u-/, in which case the nasal initial *ŋ- was simply lost.

became voiceless and aspirated regardless of the tonal conditions. In fact, the development pattern of the Middle Chinese voiced stops and affricates has been a very crucial feature in determining the classification and subgrouping of modern Chinese dialects (Li 1973, Norman 1988).

The Dapeng dialect shares this pattern with Hakka. As shown in Correspondence (1), for instance, 步 "step" and 捕 "to catch" are from the Middle Chinese Bìng 並 initial category, and 道 "road" and 稻 "paddy" are from the Dìng 定 category.[34] They are all reconstructed as having a voiced initial, which was devoiced and became aspirated in both Dapeng and Hakka (vis-à-vis unaspirated in Cantonese). In other words, although Dapeng, Hakka, and Cantonese all have both the aspirated and unaspirated voiceless obstruents in their initials inventories today, these may have different origins.

(1) Development Pattern of Middle Chinese Obstruents Initials

Dapeng	Hakka	Cantonese	MC initial	Morpheme
p^hu^{54}	p^hu^{52}	pou^{22}	*b-	步 "step"
p^hu^{35}	p^hu^{31}	pou^{22}	*b-	捕 "to catch"
t^hau^{54}	t^hau^{52}	tou^{22}	*d-	道 "road"
t^hau^{22}	t^hau^{44}	tou^{22}	*d-	稻 "paddy"
pu^{22}	pu^{52}	pou^{33}	*p-	布 "cloth"
tau^{22}	tau^{52}	tou^{33}	*t-	到 "to arrive"
p^hu^{35}	p^hu^{31}	p^hou^{35}	*p^h-	普 "normal"
t^hu^{35}	t^hu^{31}	t^hou^{35}	*t^h-	土 "earth"

---

[34] The other four morphemes, 布 "cloth", 普 "normal", 到 "to arrive", and 土 "earth" are from either the unaspirated or aspirated voiceless initial categories in Middle Chinese. They are included for the sake of parallel comparison, which highlights the notable distinction between Dapeng/Hakka and Cantonese in terms of Middle Chinese voiced obstruent development.

3.4.2 Finals

The Dapeng dialect preserves all of the six consonantal endings, both nasals (/-m/, /-n/, and /- ŋ/) and the unreleased stops (/-p/, /-t/, and /-k/), from Middle Chinese. The former set belongs to the Middle Chinese *Yángshēngyùn* 陽聲韻 rhyme group (*-m, *-n, and *-ŋ), the latter to the *Rùshēngyùn* 入聲韻 group (*-p, *-t, and *-k). The preservation of both sets of ending is an important feature that distinguishes Hakka and Cantonese from other major groups of Chinese dialects, [35] including Putonghua, in which only two of the nasal endings, /-n/ and /- ŋ/, are still kept.

The Dapeng final inventory also features the syllabic nasal /m̩/, that is, the nasal consonant can stand as legitimate syllables (and hence morphemes) in their own right. All such morphemes are from the Middle Chinese Yù 遇 final group. Both Cantonese and Hakka have syllabic nasals, the former with /m̩/ and /ŋ̍/ and the latter with /ŋ̍/ or /n̩/,[36] respectively, while none of these syllabic nasals are allowed in the regular morphemes in Putonghua.[37] All of the syllabic nasals /m̩/ in the Dapeng dialect have the glide /w-/ initial in their Putonghua counterparts.

In the meantime, there are also some finals present in Putonghua but not in the Dapeng dialect, for instance, the high front rounded vowel /y/. In fact, the phoneme /y/ does not exist in the Dapeng dialect. This feature highly resembles Hakka, while

---

[35] Therefore, Ho (2015) refers to these two varieties as the slowest in the evolution of ancient tones.
[36] Li and Zhang (1992) records mostly dental syllabic nasal [n̩]; however, according to both Beijing Daxue (2003), [ŋ̍]/[n̩] are in fact a pair of allophones.
[37] In Putonghua syllabic nasals can only be used as interjections.

Cantonese, like Putonghua, also has /y/. Example (2) below shows the cross-dialectal comparison.

(2)
Dapeng	Hakka	Cantonese	Morpheme
tsi^{42}	tsu^{44}	tsy^{53}	豬 "pig"
hɐng^{31}	hian^{11}	jyn^{21}	玄 "mysterious"
kʰi^{22}	kʰi^{44}	køy^{22}	巨 "huge"

Another aspect that makes the Dapeng dialect final inventory special is its choice of medial. As mentioned in §3.2.3, the Dapeng dialect does not allow the high back rounded vowel /u/ to appear as a medial, that is, between the initial and main vowel. Besides being a main vowel, /u/ can only occur as an ending in Dapeng. In contrast, the medial /-u-/ is common in both Putonghua and both of Dapeng's input dialects, Cantonese and Hakka. Example (3) demonstrates this distinction among Dapeng, Cantonese, and Hakka.

(3)
Dapeng	Hakka	Cantonese	Morpheme
kai^{22}	kuai^{52}	kuai^{33}	怪 "weird"
kun^{22}	kuan^{52}	kuan^{33}	慣 "habit"
kɔng^{42}	kuɔng^{44}	kuɔng^{53}	光 "light (noun)"

In the Dapeng dialect, /i/ can occur at the medial, main vowel, or ending position. This is the same as in Putonghua and Hakka. (Though in Cantonese, the medial /-i-/ is not allowed.) However, compared with both Putonghua and Hakka, the medial /-i-/ in Dapeng is considerably less frequently used. Example (4) shows the contrast, in which

Hakka has the the medial /-i-/ for all five morphemes whereas Cantonese forbids it altogether; the Dapeng dialect only uses /-i-/ in the one of the morphemes.

(4)

Dapeng	Hakka	Cantonese	Morpheme
hɐng^{31}	hian^{11}	jyn^{21}	玄 "mysterious"
kan^{42}	kian^{44}	kan^{53}	奸 "treacherous"
luk^{42}	liuk^{5}	lʊk^{2}	綠 "green"
hɐŋ^{42}	hiuŋ^{44}	hiŋ^{53}	兄 "older brother"
tsiɛ^{35}	tsia^{31}	tsɛ^{35}	姐 "older sister"

3.4.3   Tones

The Dapeng dialect, like many others in Southern China, preserves the ancient *Rù* tone 入聲 (the "entering" tone) from Middle Chinese. As indicated in §3.2.4, the *Rù* tone, which literally means the "entering" or the "checking" tone, refers to a tone that is typically shorter than others. The *Rù* tone consistently cooccurs with syllables with a stop consonant. In Dapeng, Cantonese, and Hakka, the *Rù* tone occurs with syllables ending with /-p/, /-t/, or /-k/, which are from the Middle Chinese *Rùshēngyùn* 入聲韻, literally "the rhyme group with a *Rù* tone", i.e. syllables with a stop ending (*-p, *-t, or *-k). All three stop endings, however, have been lost in Putonghua, as well as in many other Northern Mandarin dialects.

In many Cantonese dialects, however, the ancient *Rù* tone has evolved into three or four subcategories, while in the Dapeng dialect (as well as in Hakka) it only split into two, with one called the *Yīn* 陰 register and the other the *Yáng* 陽 register. Phonetically, the *Rù* tone registers do not translate equivalently across dialects as the Cantonese *Yīn* 陰

65

register(s) of the *Rù* tone normally has a higher pitch than the *Yáng* 陽 register(s), whereas in both Dapeng and Hakka the *Yīn* 陰 register has a lower pitch, as shown in Table 5.

Dapeng	Hakka	Cantonese	Register
42	1	5	*Yīn* 陰
		3	
54	5	2	*Yáng* 陽

Table 5. The Pitch Value of the *Rù* Tones

Additionally, the other three tones of Middle Chinese (*Píng* 平聲, *Shǎng* 上聲, and *Qù* 去聲) have also split into two (occasionally three) subcategories in Cantonese, that is, the *Yīn* 陰 and the *Yáng* 陽 registers respectively. The tonal split however, is not as regular in either Hakka or Dapeng: the former does not separate the *Shǎng* tone 上聲 or the *Qù* tone 去聲, and the latter does not separate the *Shǎng* tone 上聲 into the *Yīn* 陰 and the *Yáng* 陽 registers, which could both suggest some later mergers of tones.[38]

3.5   Summary

A detailed investigation of the Dapeng sound system in this chapter (as well as in Appendix C) shows some degree of resemblance between the Dapeng dialect and the source dialects—both Cantonese and Hakka. The resemblance to either source dialect is

---

[38] In the Dapeng dialect part of the ancient *Shǎng* tone 上聲 that associates with voiced and nasal initials, which are referred to as *Quánzhuó* 全濁 and *Cìzhuó* 次濁 in the traditional terminology, respectively, has merged with other tones. The merger of tones will be returned to with more details in §5.4.2.

shown at all aspects: initials, finals, and tones. This chapter focuses on the description of

the similarities as well as differences. In Chapter 5, which discusses the formation

processes of the Dapeng dialect, many of the distinctive features presented in this chapter

of this dialect will be revisited and analyzed in greater depth.

Chapter 4   Lexicon and Syntax

This chapter describes the basic lexicon and syntax of the Dapeng dialect, both being important parts of the conventional "dialect report." It first introduces the background information of the local informants. New data collection materials and procedures, which are both different from those used to collect phonetic and phonological data, are also introduced (§4.1). In §4.2, a list of basic Dapeng vocabulary based on part of speech and semantic domains is included, followed by several notes that discuss some important features of the Dapeng lexicon. Then §4.3 summarizes some distinctive syntactic features of the Dapeng dialect, which were elicited primarily from speaking tasks.

It should also be mentioned before these sections start, that morphology is not specifically addressed in this chapter. Defining wordhood in Chinese is difficult (Chao 1968); since syntactic relationships found in phrases are also all seen in compound words and given the lack of phonological cues to distinguish words and phrases, Norman (1988) has pointed out the difficulty in drawing a solid boundary between morphology and syntax in Chinese. As a typical isolating language, Chinese[39] overall lacks grammatical morphology (Li and Thompson 1981: 11) and relies heavily on word order, particles and

---

[39] Here the term "Chinese" is used as a cover term which includes Mandarin, Cantonese, Hakka, and other Sinitic languages.

prepositions to carry most of the burden of grammar (Norman 1988: 159). Therefore, instead of tackling the morphology of Dapeng in a separate section, it is discussed in both §4.2 (lexicon, for instance, the female gender marker of animals) and §4.3 (syntax, for instance, aspect markers) whenever appropriate.

4.1   Background: Informants, Materials, and Procedures

In this chapter, Mr. L still serves as an informant, but only in the grammatical survey part. Considering the heavy burden of the reading, speaking, and spontaneous speech tasks, I decided not to ask Mr. L to complete the lexical survey, which normally requires another 2-3 hours of intense investigation. Instead, I consulted Mr. W, another senior Dapeng speaker, specifically for the basic vocabularies.

Mr. W was 78 years old at the time of recording. He finished the fifth year in elementary school, then did some part-time jobs for some years, and later on worked for the local government until his 30's. After being persecuted for some political reasons in the 1960's, Mr. W quit his government job and chose farming until retirement. Except for four years in Hong Kong, from 24 to 28 years old, he had spent his entire life in Dapeng. Mr. W used the Dapeng dialect every day, both at home with other family members and in most situations in the local community. His wife is also a native speaker of the Dapeng dialect, and the people whom Mr. W interacted with are mostly other senior Dapeng speakers.

In addition to the Dapeng dialect, Mr. W claimed to also be able to speak Cantonese, Hakka, and Putonghua, with both proficiency and frequency of use of the

three major dialects ranking as Cantonese > Hakka > Putonghua. He only used these varieties of Chinese when his interlocutors were unable to understand the Dapeng dialect (for instance, with his tenants, mostly migrant workers from other provinces). Mr. W watched TV programs in both Cantonese and Putonghua, and he said he had no preference in choosing one language over the other.

The recordings with Mr. W were conducted in a quiet office within the local senior activity center. Unlike recordings with Mr. L, none of the reading, speaking, and spontaneous speech sessions were conducted; rather, the only focus was on colloquial vocabulary. The vocabulary items from the *Linguistic Atlas of Chinese Dialects* word list (Cao 2008) were elicited by picture naming (for some concrete concepts such as nouns), acting (for non-objects such as verbs and adjectives), or situation explanation (for more abstract concepts and words that are hard to act out), in order to avoid influencing the informants with the standard or written forms of lexical items. These vocabulary items were all recorded using a Tascam DR-07 MK II recorder at 44,100 Hz, in the 24-bit WAV format. The lavalier microphone was again used.

In order to describe grammatical patterns, I relied on recordings with all Dapeng speakers who participated in my fieldwork data collection. In particular, their "speaking tasks" results served as the primary source. As discussed in §2.2.2 and §2.3, grammatical patterns of the Dapeng dialect were obtained from spontaneous speech (as opposed to speech production elicited from the direct translation of a list of "sample sentences").

Three types of spoken data were recorded: story-telling, spontaneous narrations, and casual conversation. Story-telling was elicited using the picture book *Frog, Where*

are You? (Mayer 2003), the short silent *The Pear Film* (Chafe 1980, duration: approximately 6 minutes), and the famous Aesop parable, *The North Wind and the Sun*. Spontaneous narrations included several topics, such as the Spring Festival celebration, TV programs, and Dapeng history. Casual conversations were recorded mostly among local elders in the senior activity center.

While the first two types of spoken data were recorded using the Tascam DR-07 MK II recorder, the third type was recorded using an Etekcity digital audio pen recorder in the 128K bps WAV format. Since precise phonetic transcription was not the focus of the recording of casual conversations, the pen recorder was favored due to its mobility. More importantly, a large pencil-box sized recording device, such as the Tascam recorder was somewhat incompatible with the casual nature of the conversations. Native speakers' consciousness clearly rose as soon as they saw the recorder, and the distance between the interviewer and the local community immediately increased. A pen recorder, on the other hand, was much less conspicuous, and therefore the Dapeng residents spoke more naturally in front of it.

All speaking tasks were transcribed by Miss Y, a native 22-year old speaker of Dapeng who was majoring in English at a nearby college. A representative transcription of Mr. L telling the *Frog, Where Are You?* story is included in Appendix B. Translation in both Standard Chinese and English are also provided.

4.2   Lexicon: A Basic Vocabulary List

This section provides a list of the basic vocabulary of the Dapeng dialect. All lexical items are grouped based on semantic domains, in which a set of words are related and share certain similarities in meaning. These lexical items listed are primarily from the *Linguistic Atlas of Chinese Dialects* word list (Cao 2008) and are also supplemented by a number of words collected from the transcription of Dapeng speakers' speaking tasks as well as their casual conversations.

This vocabulary list includes both cognates, which the Dapeng dialect shares with adjacent major varieties of Chinese (Cantonese and/or Hakka) or with Chinese dialects in general, and special, colloquial lexical items that only exist in the Dapeng dialect. In the latter category, an attempt was also made to identify the original morphemes (考本字 *kǎo běnzì*) in the Dapeng dialect which are, again, orthographically represented by Chinese characters.

In this chapter, the identification of original morphemes is based on two criteria proposed by Mei (1995) and Yang (2000), i.e. phonological correspondence and semantic similarity. That is, the pronunciation of the proposed morphemes should show strict phonological correspondence in line with the overall diachronic sound change of other morphemes in the same dialect; concurrently, the meaning of the proposed morphemes should match the meaning of the colloquial words. In addition, the native speakers' intuition is also consulted and considered.

It should also be noted that identifiable morphemes are not always readily available. This is especially the case for Southern Chinese dialects, which are known for

having non-Sinitic substrata due to long-term contact with non-Han ethnic groups. Original morphemes of loanwords from those non-Han languages are not always unidentifiable. Morphemes with no character detectable will be represented by the symbol □.[40]

In total, all investigated Dapeng vocabulary items are divided into five groups based on part of speech, each including tens of lexical items. For each item, the following information is provided: glosses in English and in Standard Chinese, pronunciation (in IPA), and original morphemes (in characters, if identifiable and if different from Standard Chinese).

4.2.1 Nouns

Nature

1.	'sun' 太陽	ŋit^{54} tʰɐu^{31}	熱頭
2.	'moon' 月亮	it^{54} kɔŋ^{42}	月光
3.	'thunder' 雷	(lui^{31}) kuŋ^{42} hiɔŋ^{35}	（雷）公響
4.	'flash' 電	fɔ^{35} sa^{31} tsɐi^{35}	火蛇仔
5.	'rain' (as a verb) 下雨	lɔk^{54} sui^{35}	落水
6.	'rainbow' 虹	tin^{42} kuŋ^{42}	天□
7.	'water pit' 水坑兒	sui^{35} tʰɐm^{31}	水氹

Everyday life

8.	'village' 村莊	tsʰin^{42}	村
9.	'alley' 胡同	hɔŋ^{54}	巷
10.	'home' 家	uk^{42} kʰi^{22}	屋企
11.	'cement' 水泥	sui^{35} nɐi^{31}	

---

[40] The character-less morphemes is referred to in the traditional terminology as a "sound (morpheme) without identifiable character" (*yǒu shēng wú zì* 有聲無字).

12.	'house' 房子	uk^{42}	屋
13.	'apartment' (multi-floors) 樓	lɐu^{31}	
14.	'room' 屋子	fɔŋ^{31}	房
15.	'bedroom' 臥室	fɔŋ^{31}	房
16.	'window' 窗	tsʰɔŋ^{42}	
17.	'threshold' 門檻兒	mun^{31} tam^{42}	門口
18.	'oven' 灶	tsau^{22} tʰɐu^{31}	灶頭
19.	'stove' 爐	lu^{31} tsɐi^{35}	爐仔
20.	'kitchen god' 灶神	tsau^{22} tʰɐu^{31} kuŋ^{42}	灶頭公
21.	'pot' 鍋	wɔk^{54}	鑊
22.	'kitchen knife' 刀	tau^{42}	
23.	'firewood' 柴	tsʰai^{31}	
24.	'(old style) toilet' 廁所	si^{35} kɔŋ^{42}	屎缸
25.	'pigsty' 豬圈	tsi^{42} tsʰau^{31}	豬巢
26.	'nestle' 鳥窩	tsiɔk^{42} tau^{22}	雀竇
27.	'bed' 床	min^{31} tsʰɔŋ^{31}	眠床
28.	'quilt' 被子	min^{31} tʰui^{42}	棉胎
29.	'table' 桌子	tʰui^{35}	臺
30.	'drawer' 抽屜	tʰɔ^{42} siɔŋ^{42}	拖箱
31.	'bowl' 碗	un^{35}	
32.	'claypot' 瓦煲	pau^{42}	煲
33.	'chopsticks' 筷子	tsʰi^{54}	箸
34.	'spoon' 湯匙	tʰiu^{31} kaŋ^{42}	調羹
35.	'bottle' 瓶子	aŋ^{42}	罌
36.	'lid' 蓋子	kui^{22}	蓋
37.	'kitchen' 廚房	lɔŋ^{31} tsɐi^{35}	廊仔
38.	'bicycle' 自行車	tan^{42} tsʰa^{42}	單車
39.	'wheel' 輪子	lɐn^{31}	輪
40.	'umbrella' 雨傘	tsa^{42}	遮
41.	'clothes' 衣服	sam^{42} fu^{22}	衫褲
42.	'shoelace' 鞋帶	hai^{31} sin^{31}	鞋繩
43.	'diaper' 尿布	niu^{54} pʰin^{35}	尿片
44.	'pocket' 口袋	tʰui^{54}	袋
45.	'sleeves' 袖子	sam^{42} tsʰɐu^{54}	衫袖
46.	'towel' 毛巾	sɐu^{35} kɐn^{42}	手巾
47.	'soap' 肥皂	fan^{42} kan^{35}	番城
48.	'hot water' 熱水	ŋit^{54} sui^{35}	熱水
49.	'boiling water' 沸水	kʰɐn^{35} sui^{35}	滾水
50.	'comb' 梳子	sɔ^{42}	梳

51.	'scissors' 剪刀	kau⁴² tsʰin³⁵	鉸剪
52.	'hoe' 鋤頭	kiɔk⁴² tsʰɔ³¹	腳鋤
53.	'grave' 墳墓	fɐn³¹ tɐu³¹	墳頭
54.	'fire (disaster)' 火災	fɔ³⁵ tsuk⁴²	火燭

People[41]

55.	'human' 人	jɐn³¹	
56.	'guest' 客人	hak⁴² jɐn³¹	
57.	'married woman' 已婚女人	fu⁵⁴ niɔŋ³¹	婦娘
58.	'child' 小孩	sɐi²² mɐn⁴² tsɐi³⁵	細蚊仔
59.	'boy' 男孩	nam³¹ tsɐi³⁵	男仔
60.	'girl' 女孩	ni³⁵ tsɐi³⁵	女仔
61.	'young adult (male)' 年輕男性	tsɐi³⁵	仔
62.	'young adult (female)' 年輕女性	ni³⁵	女
63.	'bagger' 乞丐	lɔ³¹ sik⁵⁴ lau³⁵	攞食佬
64.	'grandfather' (paternal) 祖父	a²² kuŋ⁴²	阿公
65.	'grandmother' (paternal) 祖母	a²² pʰɔ³¹	阿婆
66.	'grandfather' (maternal) 外祖父	tsiɛ³⁵ kuŋ⁴²	姐公
67.	'grandmother' (maternal) 外祖母	tsiɛ³⁵ pʰɔ³¹	姐婆
68.	'father' 父親	a²² pa⁴²	阿爸
69.	'mother' 母親	a²² mi⁴²	阿嬤
70.	'husband's father' 公公	ka⁴² kuŋ⁴²	家公
71.	'husband's mother' 婆婆	ka⁴² pʰɔ³¹	家婆
72.	'father's older brother' 伯父	a²² pak⁴²	阿伯
73.	'father's older brother's wife' 伯母	pak⁴² niɔŋ³¹	伯娘
74.	'father's younger brother'叔叔	a²² suk⁴²	阿叔
75.	'father's younger brother's wife' 叔母	a²² sɐm³⁵	阿嬸
76.	'father's older sister' 大姑	tʰai⁵⁴ ku⁴²	大姑
77.	'father's younger sister' 小姑	ku⁴² tsɐi³⁵	姑仔
78.	'mother's older brother' 舅舅	tʰai⁵⁴ kʰɐu²²	大舅
79.	'mother's younger brother' 舅舅	kʰɐu²² tsɐi³⁵	舅仔
80.	'mother's brother's wife' 舅母	a²² kʰɐm⁴²	阿妗
81.	'mother's older sister' 大姨	tʰai⁵⁴ i³¹	
82.	'mother's younger sister' 小姨	i³¹ tsɐi³⁵	姨仔

[41] Kinship terms in this section are primarily terms of address, not necessarily terms of reference. During the fieldwork the kinship terms were elicited by the question "how would you address your [relative A]'s [relative B]." For instance, the informants were asked about the way to address their paternal grandmother by the question "how would you address your father's mother."

83. 'elder brother' 哥哥 a²² kɔ⁴² 阿哥
84. 'elder sister' 姐姐 a²² tsiɛ³⁵ 阿姐
85. 'younger brother' 弟弟 a²² tʰɐi⁵⁴ 阿弟
86. 'younger sister' 妹妹 a²² mui⁵⁴ 阿妹
87. 'son' 兒子 a²² tsɐi³⁵ 阿仔
88. 'daughter-in-law' 媳婦 sɐn⁴² pu³⁵ 新婦
89. 'daughter' 女兒 a²² ni³⁵ 阿女
90. 'son-in-law' 女婿 ni³⁵ sɐi²² 女婿
91. 'grandchild' (son's side) 孫 sin⁴²
92. 'grandchild' (daughter's side) 外孫 ŋui⁵⁴ sin⁴²
93. 'nephew' (brother's son) 侄 tsʰɐt⁵⁴ tsɐi³⁵ 侄仔
94. 'nephew' (sister's son, speaker is male) 外甥 ŋui⁵⁴ saŋ⁴² 外生
95. 'husband' 丈夫 lau³⁵ kuŋ⁴² 老公
96. 'wife' 妻子 lau³⁵ pʰɔ³¹ 老婆
97. 'bride' 新娘 sɐn⁴² niɔŋ³¹

Parts of the body

98. 'head' 頭 tʰɐu³¹
99. 'hair' 頭髮 tʰɐu³¹ mau³¹ 頭毛
100. 'braid 辮子 pin⁴² 辮
101. 'face' 臉 min⁵⁴ 面
102. 'eye' 眼 ŋan²²
103. 'tear' 眼淚 ŋan²² lui⁵⁴
104. 'nose' 鼻子 pʰi⁵⁴ kuŋ⁴² 鼻公
105. 'nasal mucus'(thin) 清鼻涕 pʰi⁵⁴ sui³⁵ 鼻水
106. 'nasal mucus'(thick) 濃鼻涕 pʰi⁵⁴ nuŋ³¹ 鼻膿
107. 'ear' 耳朵 ŋi³⁵ kɐt⁵⁴ 耳□
108. 'mouth' 嘴巴 tsui³⁵ 嘴
109. 'tooth' 牙齒 ŋa³¹ 牙
110. 'tongue' 舌頭 li⁵⁴ 脷
111. 'saliva' 口水 hɐu³⁵ sui³⁵ 口水
112. 'moustache' 鬍子 su⁴² 鬚
113. 'neck' 脖子 kiaŋ³⁵ 頸
114. 'throat' 喉嚨 hɐu³¹ luŋ³¹
115. 'hand' 手 sɐu³⁵
116. 'arm' 臂 sɐu³⁵ 手
117. 'left hand' 左手 tsɔ³⁵ sɐu³⁵ 左手

118. 'right hand' 右手   jɐu⁵⁴ sɐu³⁵   右手
119. 'fist' 拳頭   kʰin³¹   拳
120. 'finger' 手指頭   sɐu³⁵ tsi³⁵   手指
121. 'nail' 手指甲   sɐu³⁵ tsi³⁵ kap⁴²
122. 'shoulder' 肩膀   kin⁴² tʰɐu³¹   肩頭
123. 'underarms' 腋下   lɐk⁵⁴ tsʰak⁵⁴ ha⁵⁴   肋赤下
124. 'foot' 腳   kiɔk⁴²
125. 'leg' 腿   kiɔk⁴²   腳
126. 'knee' 膝蓋   sɐt⁴² tʰɐu³¹   膝頭
127. 'belly' 肚子   tu³⁵ pʰat⁵⁴   肚□
128. 'back' 背   tsʰɐk²² lau³¹   脊□
129. 'buttock' 屁股   si³⁵ fɐt⁴² tʰɐn³¹   屎□□
130. 'anus' 肛門   si³⁵ fɐt⁴² ŋan³⁵   屎□眼
131. 'breast' 乳房   nin⁴²   □
132. 'penis' 陰莖   lɐn³⁵   □
133. 'vagina' 女陰   hɐi⁴²   屄

Meal and food

134. 'congee' 稀飯   tsuk⁴²   粥
135. 'steamed bun' 饅頭   man³¹ tʰɐu³¹
136. 'steamed stuffed bun' 包子   pau⁴²   包
137. 'deep-fried twisted dough sticks' 油條   jɐu³¹ tsa⁵⁴ kɐi³⁵   油炸鬼
138. 'dish/course' 菜   suŋ²²   餸
139. 'pig liver' 豬肝   tsi⁴² kun⁴²
140. 'pig tongue' 豬舌頭   tsi⁴² li⁵⁴   豬脷
141. 'salt' 鹽   im³¹
142. 'vinegar' 醋   tsu²²
143. 'burnt rice' 鍋巴   fan⁵⁴ tsiu⁴²   飯焦
144. 'leftover (meal)' 剩菜   tsʰui²² kiɔk⁴²   菜腳

Animals

145. 'male pig' 公豬   tsi⁴² kuŋ⁴²   豬公
146. 'male pig' (young) 小公豬   tsi⁴² ku³⁵   豬牯
147. 'female pig' 母豬   tsi⁴² pʰɔ³¹   豬婆
148. 'male dog' 公狗   kɐu³⁵ ku³⁵   狗牯
149. 'female dog' 母狗   kɐu³⁵ na³⁵   狗嫲

150.	'rooster' 公雞	kɐi⁴² kuŋ⁴²	雞公
151.	'hen' 母雞	kɐi⁴² na³⁵	雞乸
152.	'chick' 小雞	kɐi⁴² tsɐi³⁵	雞仔
153.	'(chicken) egg' 雞蛋	kɐi⁴² tsʰɐn⁴²	雞春
154.	'bird' 鳥兒	tsiɔk⁴² tsɐi³⁵	雀仔
155.	'sparrow' 麻雀	ma³¹ tsiɔk⁴² tsɐi³⁵	麻雀仔
156.	'butterfly' 蝴蝶	fu21 tʰip⁵⁴	
157.	'dragonfly' 蜻蜓	lɔŋ³¹ ni 21	蜋口
158.	'mouse' 老鼠	lau³⁵ si³⁵	
159.	'bat' 蝙蝠	mɐn⁴² si³⁵	蚊死
160.	'tiger' 老虎	lau³⁵ fu³⁵	
161.	'cat' 貓	miu³⁵	
162.	'monkey' 猴子	hɐu³¹ tsɐi³⁵	猴仔
163.	'snake' 蛇	sa³¹	
164.	'earthworm' 蚯蚓	sa³¹ hin³⁵ (kuŋ⁴²)	蛇蜆（公）
165.	'spider' 蜘蛛	kʰam³¹ lau³¹	蟲蟧
166.	'ant' 螞蟻	ŋɐi³⁵	
167.	'mosquito' 蚊子	mɐn⁴²	蚊
168.	'fly' 蒼蠅	u⁴² jɐŋ³¹	烏蠅
169.	'flea' 跳蚤	kɐu³⁵ sɐt⁴²	狗虱
170.	'cockroach' 蟑螂	kʰit⁴² tsʰat⁴²	甴曱
171.	'firefly' 螢火蟲	fɔ³⁵ im³¹ tsʰuŋ³¹	火炎蟲
172.	'frog' (bigger ones) 大青蛙	kap⁵⁴ na³⁵	蛤乸
173.	'frog' (smaller ones) 小青蛙	kɐi³⁵ tsɐi³⁵	蜎仔
174.	'toad' 蟾蜍	kɐm³¹ si³¹	蟾蜍
175.	'scale (of fish)' 鱗	lɐŋ³¹	
176.	'wings (of bird)' 翅膀	jɐk⁵⁴	翼

Plants and vegetables

177.	'rice plant' 稻	wo³¹	禾
178.	'rice seed' 稻穀	kuk⁴²	穀
179.	'rice straw' 稻草	wo³¹ kun³⁵ tsʰau³⁵	禾管草
180.	'flour' 麵粉	min⁵⁴ fɐn³⁵	
181.	'millet' 穀子	kɐu³⁵ mi²² suk⁴²	狗尾粟
182.	'corn' 玉米	pau⁴² suk⁴²	包粟
183.	'powder' 面兒	fɐn³⁵	粉
184.	'horsebean' 蠶豆	tsʰam³¹ tʰɐu⁵⁴	

185. 'pea' 豌豆	tʰɐu⁵⁴ mɐi³⁵	豆米
186. 'peanut' 花生	tʰi⁵⁴ tʰɐu⁵⁴	地豆
187. 'sunflower' 向日葵	kʰɐi³¹ fa³⁵	葵花
188. 'radish' 蘿蔔	lɔ³¹ pʰak⁵⁴	
189. 'spinach' 菠菜	kɔk⁴² tsui²²	角菜
190. 'cabbage' 圓白菜	jai³¹ tsui²²	椰菜
191. 'eggplant' 茄子	kʰiɔ³¹ tsɐi³⁵	茄仔
192. 'hot pepper' 辣椒	lat⁵⁴ tsiu⁴²	辣椒
193. 'mushroom' 蘑菇	mɔ³¹ ku⁴²	
194. 'sweet potato' 甘薯	fan⁴² si³¹	番薯
195. 'potato' 馬鈴薯	si³¹ sɐi³⁵	薯仔
196. 'tomato' 西紅柿	fan⁴² kʰiɔ³¹	番茄
197. 'pumpkin' 南瓜	kɐn⁴² ka⁴²	金瓜
198. 'towel gourd' 絲瓜	sui³⁵ ka⁴²	水瓜
199. 'fruit' 水果	saŋ⁴² kɔ³⁵	生果

Time and space

200. 'this year' 今年	kɐm⁴² nin³¹	
201. 'next year' 明年	mɐŋ³¹ nin³¹	
202. 'the year after next year' 後年	hɐu⁵⁴ nin³¹	
203. 'last year' 去年	kʰɐu⁵⁴ nin³¹	舊年
204. 'the year before last year' 前年	tsʰin³¹ nin³¹	前年
205. 'today' 今天	kɐm⁴² jɐt⁴²	今日
206. 'tomorrow' 明天	mɐŋ³¹ jɐt⁴²	明日
207. 'the day after tomorrow' 後天	hɐu⁵⁴ jɐt⁴²	後日
208. 'two days after tomorrow' 大後天	tʰai⁵⁴ hɐu⁵⁴ jɐt⁴²	大後日
209. 'yesterday' 昨天	tsʰɔŋ⁵⁴ jɐt⁴²	口日
210. 'the day before yesterday' 前天	tsʰin³¹ jɐt⁴²	前日
211. 'two days before yesterday' 大前天	tʰai⁵⁴ tsʰin³¹ jɐt⁴²	大前日
212. 'daytime' 白天	jɐt⁴² tʰau³¹	日頭
213. 'nighttime' 晚上	ia⁵⁴ tʰau³¹	夜頭
214. 'early morning' (until 8am) 早上	tsʰaŋ⁴² tsau³⁵	清早
215. 'morning' (before 11am) 上午	sɔŋ⁵⁴ tsɐu²²	上晝
216. 'noon' (around 12pm) 中午	an²² tsɐu²²	晏晝
217. 'afternoon' (2pm – 5pm) 下午	ha⁵⁴ tsɐu²²	下晝
218. 'evening' 傍晚	man²² hɐk⁴²	晚黑
219. 'night' 晚上	man²² hɐk⁴²	晚黑

220. 'above' 上面   sɔŋ⁵⁴ kau⁴²   上高
221. 'below' 下面   ha⁵⁴ tɐi⁴²   下低
222. 'front' 前面   tsʰin³¹ tʰau³¹   前頭
223. 'back' 後面   hɐu⁵⁴ mi²²   後尾
224. 'inside' 裡面   nui⁵⁴ tʰau³¹   內頭
225. 'outside' 外面   ŋui⁵⁴ min⁵⁴   外面

4.2.2 Pronouns, Numbers, and Classifiers (CL)

226. 'first person, singular' 我   ŋɔ³⁵
227. 'second person, singular' 你   ni²²
228. 'third person, singular' 他   kʰi²²   佢
229. 'first person, plural' (listener-exclusive) 我們   ŋɔ³⁵ tʰi⁵⁴   我哋
230. 'first person, plural' (listener-inclusive) 咱們   ŋɔ³⁵ tʰi⁵⁴   我哋
231. 'second person, plural' 你們   ni²² tʰi⁵⁴   你哋
232. 'third person, plural' 他們   kʰi²² tʰi⁵⁴   佢哋
233. 'everyone' (all-inclusive) 大家   tʰai⁵⁴ ka⁴²   大家
234. 'self' 自己   tsʰi⁵⁴ ki²²
235. 'other people' 別人   jɐn³¹ tʰi⁵⁴   人哋
236. 'proximal demonstrative' 這   ni³⁵   呢
237. 'distal demonstrative' 那   kɔ³⁵   嗰
238. 'this one' 這個   ni³⁵ tsik⁵⁴   呢隻
239. 'that one' 那個   kɔ³⁵ tsik⁵⁴   嗰隻
240. 'which one' 哪個   na³⁵ tsik⁵⁴   哪隻
241. 'who' 誰   na⁵⁴ jɐn³¹   哪人
242. 'here' 這裡   ni³⁵ tʰiaŋ⁴²   呢廳
243. 'there' 那裡   kɔ³⁵ tʰiaŋ⁴²   嗰廳
244. 'where' 哪裡   na³⁵ tʰiaŋ⁴²   哪廳
245. 'so' 這麼   kam²²   咁
246. 'how' 怎麼   tim³⁵ (iɔŋ⁵⁴)   點（樣）
247. 'how many/much' 多少   ki³⁵ tɔ⁴²   幾多
248. 'what' 什麼   mɐt⁴² (ja³⁵)   乜（嘢）
249. 'do what' 幹什麼   tsu²² mɐt⁴² (ia³⁵)   做乜（嘢）
250. 'why' 為什麼   tim³⁵ kai³⁵   點解
251. 'thing/object' (in general) 東西   ia³⁵   嘢
252. 'thing/matter (in general) 事情   si⁵⁴   事

253. 'one' 一 jɐt^{42}

254. 'two' 兩 liɔŋ^{35}

二 ŋi^{54}

255. 'three' 三 sam^{42}

256. 'four' 四 si^{22}

257. 'five' 五 m̩^{35}

258. 'six' 六 luk^{54}

259. 'seven' 七 tsʰɐt^{42}

260. 'eight' 八 pat^{42}

261. 'night' 九 kɐu^{35}

262. 'ten' 十 sɐp^{54}

263. 'eleven' 十一 $\text{sɐp}^{54}\,\text{jɐt}^{42}$

264. 'twelve' 十二 $\text{sɐp}^{54}\,\text{ŋi}^{54}$

265. 'thirteen' 十三 $\text{sɐp}^{54}\,\text{sam}^{42}$

266. 'twenty' 二十 $\text{liɔŋ}^{35}\,\text{sɐp}^{54}$ 兩十

267. 'hundred' 百 pak^{42}

268. 'thousand' 千 tsʰin^{42}

269. 'ten thousand' 萬 man^{54}

270. 'CL[42] for people' 個 tsik^{54} 隻

271. 'CL for the Chinese currency" 元 mɐn^{42} 文

ŋɐn^{31} 銀

角 hau^{31} 毫

272. 'CL for people' 個（人） tsik^{54} 隻

273. 'CL for cows' 頭（牛） tʰiu^{31} 條

274. 'CL for pigs' 頭（豬） tsik^{54} 隻

275. 'CL for dogs' 隻（狗） tʰiu^{31} 條

276. 'CL for chicken'隻（雞） tsik^{54}

277. 'CL for mosquitos'隻（蚊子） lɐp^{54} 粒

278. 'CL for fish' 條（魚） tʰiu^{31}

279. 'CL for snakes' 條（蛇） tʰiu^{31}

280. 'CL for tables' 張（桌子） tsɔŋ^{42}

281. 'CL for quilts' 床（被子） tsɔŋ^{42} 張

282. 'CL for mattresses' 張（席子） tsɔŋ^{42}

283. 'CL for pairs of shoes' 雙（鞋） tui^{22} 對

284. 'CL for knives'把（刀） tsɔŋ^{42} 張

285. 'CL for locks' 把（鎖） ba^{35}

---

[42] "CL" stands for "classifier."

286. 'CL for ropes' 根（繩子）   tʰiu³¹   條
287. 'CL for pens' 支（毛筆）   tsi⁴²
288. 'CL for doors' 扇（小門）   tsik⁵⁴   隻
289. 'CL for gates' 道（大門）   fu⁵⁴   副
290. 'CL for cars' 輛（汽車）   ka²²   架
291. 'CL for bridges' 座（橋）   tʰiu³¹   條
292. 'CL for roads' 條（路）   tʰiu³¹
293. 'CL for trees' 棵（樹）   tɐu⁴²   兜
294. 'CL for beans' 粒（豆子）   lɐp⁵⁴
295. 'CL for meals' 頓（飯）   tsʰan⁴²   餐
296. 'CL for cases of matters/events' 件（事情）   kʰin⁵⁴
297. 'a handful of' (rice) 把（米）   tsa⁴²   抓
298. 'a little bit' (amount of something) 一點兒   jɐt⁴² ti⁴²   一□
299. 'some' (amount of something) 一些   jɐt⁴² ti⁴²   一□
300. 'CL for actions' 一下   jɐt⁴² ha⁵⁴
301. 'a short period of time' 一會兒   jɐt⁴² tsʰɐn⁵⁴   一陣
302. The CL by default   tsik⁵⁴   隻

4.2.3   Verbs

303. 'work' 幹活兒   tsu²² ia³⁵   做嘢
304. 'watch' (TV) 看（電視）   tʰɐi³⁵   睇
305. 'listen' 聽   tʰiaŋ⁴²
306. 'smell' 聞   mɐn³¹
307. 'bite' 咬   ŋau³⁵
308. 'chew' 嚼   tsʰiu⁵⁴
309. 'sting' 叮   tiu⁴²   叼
310. 'lick' 舔   sai²²   □
311. 'suck' 吮吸   tsit⁴²   □
312. 'spit' 吐   tʰiɔ⁴²   □
313. 'vomit' 嘔吐   au³⁵   嘔
314. 'take' 拿   kʰai²²   [扌戒]⁴³
315. 'give' 給   pi³⁵   畀
316. 'sew' 縫   pu³⁵   補
317. 'hold (an umbrella) 打傘   kʰiɛ³¹ tsa⁴²   □遮

---

⁴³ Since the orthographic representation of the morpheme "to take" is not a standard character (and is thus untypable), I separate the character by its radical/semantic component 扌 and phonetic component 戒 and put them in square brackets to indicate their combination into a character.

318. 'pinch' 掐   lak⁵⁴   □
319. 'pinch' (move upwards) 掐并向上提   lim⁵⁴   □
320. 'take on (clothes)' 穿（衣服）   tsɔk⁴²   著（衫）
321. 'take off (clothes)' 脫（衣服）   pɔk⁴²/mɔk⁴²   剝（衫）
322. 'tie (shoelaces)' 繫（鞋帶）   tʰau³¹   □（鞋繩）
323. 'break off' (with fingers and thumb) 掰   mɐk⁵⁴   □
324. 'tweak' 擰   nɐu³⁵   扭
325. 'squeeze' 捻   nɐn³⁵   □
326. 'mix' 和（麵）   nau⁵⁴   □
327. 'knead' 揉   tsʰai⁴²   攄
328. 'tear' 撕   si⁴²
329. 'bend' 折   au³⁵   拗
330. 'pull out' 拔   tsʰa³⁵   扯
331. 'stand' 站   kʰi³⁵   徛
332. 'lean on' 倚   pʰaŋ⁴²   □
333. 'squad' 蹲   mɐu⁴² /pɐu⁴²   □
334. 'jump' 跳   tʰiu²²
335. 'stride' 邁   lam²²   □
336. 'step on' 踏   tsʰai³⁵   踩
337. 'stomp' 跺   tʰam⁵⁴   扰
338. 'crawl' 爬   pʰa³¹
339. 'walk' 走   hɐŋ³¹   行
340. 'run' 跑   tsɐu³⁵   走
341. 'escape' 逃   tsɐu³⁵   走
342. 'chase' 追   tʰiak⁴²   □
343. 'catch/arrest' 抓   tsuk⁴²   捉
344. 'hold" (in arms) 抱   lam³⁵   攬
345. 'push' 推   thɔi⁴²
346. 'fall down" 摔   tʰit⁵⁴   跌
347. 'bump into' 撞   tsʰɔŋ⁵⁴
348. 'hide' 躲藏   piaŋ²²   □
349. "put" 放   fɔŋ²²
350. 'pile up' 摞   tsʰɐn⁴²   □
351. 'bury' 埋   mai³¹
352. 'cover' (with lid) 蓋   kʰam³⁵   □
353. 'press' (from above) 壓   tsat⁵⁴   室
354. 'press' (with fingers) 摁   kʰam⁵⁴   撳
355. 'stab' 捅   tʰuŋ³⁵
356. 'hack' 砍   tsam³⁵   斬

357. 'chop/mince' 剁　　　　　　tiɔk⁵⁴
358. 'hoe' 鋤　　　　　　　　　tsʰɔ³¹
359. 'peel' 削　　　　　　　　　pʰɐi⁴²　　　□
360. 'crack' 裂　　　　　　　　pau²²　　　爆
361. 'claps' 倒塌　　　　　　　lam²²　　　冧
362. 'wipe' 擦　　　　　　　　kiu³⁵　　　□
363. 'pour away' 倒　　　　　　sa³⁵　　　灑
364. 'throw away/discard' 扔　　tɐm³⁵　　　□
365. 'toss' 投　　　　　　　　mak⁵⁴　　　□
366. 'fall' 掉　　　　　　　　tit⁵⁴　　　跌
367. 'lose' 丢　　　　　　　　lai³¹　　　落
368. 'find/look for' 找　　　　tsʰɐm³¹　　尋
369. 'pick up' (from the ground) 撿　　kʰim³⁵
370. 'lift up' (using one hand) 提　　mia⁵⁴　　　　　　　□
371. 'lift up' (using both hands) 抬　　tʰui³¹
372. 'carry' (on shoulder) 挑　　tam⁴²　　　擔
373. 'choose' 挑選　　　　　　kan³⁵　　　揀
374. 'weigh' (using a scale) 稱　　tsɐŋ³¹
375. 'earn' (money) 賺　　　　tsʰan⁴²
376. 'owe' 欠　　　　　　　　tsaŋ⁴²　　　□
377. 'chat' 聊天　　　　　　　kʰɐŋ⁴² kɐi³⁵　傾偈
378. 'speak' 說話　　　　　　kɔŋ³⁵ ia³⁵/wa⁵⁴　講嘢/話
379. 'call out' (to someone) 叫　　ham²²　　　喊
380. 'call' (a name) 稱呼　　　ham²²　　　喊
381. 'be angry' 生氣　　　　　kuk⁵⁴ hi²²　焗氣
382. 'scold/curse' 罵　　　　　nau⁵⁴　　　鬧
383. 'cry' 哭　　　　　　　　huk⁴²
384. 'quarrel' (verbal) 吵架　　tsau³¹ kau⁴²　嘈交
385. 'fight' (physical) 打架　　da³⁵ kau⁴²　打交
386. 'hit' (by hand) 打　　　　da³⁵
387. '(man) marry (woman)' 娶　　tsʰui³⁵
388. 'sleep' 睡覺　　　　　　fɐn²² kau²²　訓覺
389. 'snore' 打鼾　　　　　　ta³⁵ pʰi⁵⁴ fun³¹　打鼻鼾
390. 'take a shower' 洗澡　　　tsʰuŋ⁴² liɔŋ³¹　沖涼
391. 'play' 玩兒　　　　　　liu⁵⁴　　　撩
392. 'castrate (a pig)' 閹　　　im⁴²
393. 'slaughter (a pig)' 宰殺　　tʰɔŋ⁴²　　劏
394. 'cut (the throat)' 割　　　sin⁴²　　　□
395. 'herd cows" 放牛　　　　tsɔŋ³⁵ ŋɐu³¹　掌牛

396.	'(birds) lay eggs' 下（蛋）	saŋ⁴²	生
397.	'(birds) incubate eggs' 孵（蛋）	pʰu⁴²	
398.	'transplant (rice seedlings)' 插秧	si⁴² tin³¹	蒔田
399.	'(quickly) cook (in boiling water)' 灼	lɔk⁵⁴	烙
400.	'fry' 煎	tsin⁴²	
401.	'deep fry' 炸	tsa⁵⁴	
402.	'eat' (a meal) 吃	sik⁵⁴	食
403.	'drink' (alcohol) 喝	sik⁵⁴	食
404.	'drink' (tea) 喝	sik⁵⁴	食
405.	'smoke' (cigarette) 抽	sik⁵⁴	食
406.	'pick' (using chopsticks) 夾	kap⁴²	
407.	'(re)fill' (a wine cup) 斟	tsɐm⁴²	
408.	'have sexual intercourse with' �бро冚	diu³⁵	屌
409.	'defecate' 拉屎	ɔ⁴² si³⁵	屙屎
410.	'urinate' 拉尿	ɔ⁴² niu⁵⁴	屙尿
411.	'flatulate' 放屁	da³⁵ pʰi²²	打屁
412.	'cough' 咳嗽	kʰɐt⁴²	咳
413.	'have diarrhea' 拉肚子	tu³⁵ ɔ⁴²	肚屙
414.	'die' 死	si³⁵	死
415.	'pass away' (elders)	lau³⁵/tsɐu³⁵/ko²² sɐn⁴²	老/走/過身
416.	'understand' 知道	ti⁴²	□
417.	'not understand' 不知道	m̩³¹ tsi⁴²	唔□
418.	'know' (someone) 認識	sek⁴²	識
419.	"not know" (someone) 不認識	m̩³¹ sek⁴²	唔識
420.	'think' 想	siɔŋ³⁵	
421.	'fear' 怕	kʰɔŋ³¹	□
422.	'want' 要	ui²²	□
423.	'have' 有	jɐu²²	
424.	'not have' 沒有	mau³⁵	冇
425.	'be' (copula) 是	hɐi⁵⁴	係
426.	'be not' (copula) 不是	m̩³¹ hɐi⁵⁴	唔係
427.	'be at/in/on' 在	tsʰui²²	在
428.	'be not at/in/on' 不在	m̩³¹ tsʰui²²	唔在

4.2.4　Adjectives and Adverbs

429.　'red' 紅　　　　　　　　huŋ³¹
430.　'yellow' 黃　　　　　　wɔŋ³¹
431.　'black' 黑　　　　　　hɐk⁴²
432.　'many/much' 多　　　　tɔ⁴²
433.　'little/few' 少　　　　siu³⁵
434.　'big' 大　　　　　　　tʰai⁵⁴
435.　'small' 小　　　　　　sɐi²²　　　　　　細
436.　'long' 長　　　　　　tsʰɔŋ³¹
437.　'short'(length) 短　　tin³⁵
438.　'wide' 寬　　　　　　fut⁴²　　　　　　闊
439.　'narrow' 窄　　　　　kʰip⁵⁴　　　　　狹
440.　'tall' 高　　　　　　kau⁴²
441.　'short'(height) 矮　　ai³⁵
442.　'high' 高　　　　　　kau⁴²
443.　'low' 低　　　　　　tɐi⁴²
444.　'askew' 歪　　　　　tsʰia³¹　　　　　斜
445.　'curved' 彎　　　　　wan⁴²
446.　'steep' 陡　　　　　tsʰia³¹　　　　　斜
447.　'salty' 咸　　　　　ham21
448.　'plain' (flavor) 淡　　tʰan²²
449.　'thick' (height) 厚　　pʰɐn²²　　　　　□
450.　'thin' (height) 薄　　pʰɔk⁵⁴
451.　'thick' (liquid) 稠　　kʰit⁵⁴　　　　　結
452.　'thin' (liquid) 稀　　tsʰaŋ⁴²　　　　　清
453.　'dense' 密　　　　　mɐt⁵⁴
454.　'sparse' 稀　　　　　sɔ⁴²　　　　　　疏
455.　'bright' 亮　　　　　kɔŋ⁴²　　　　　光
456.　'dark' 黑　　　　　　hɐk⁴²
457.　'hot' (temperature) 熱　ŋit⁵⁴
458.　'cold' (temperature) 冷　laŋ³⁵
459.　'dry' 乾　　　　　　tsau⁵⁴　　　　　燥
460.　'dry' (pond) 乾　　　lim³⁵　　　　　　漧
461.　'wet' 濕　　　　　　sɐp⁴²
462.　'clean' 乾淨　　　　tsʰaŋ⁵⁴　　　　　淨
463.　'dirty' 髒　　　　　ŋɐn³⁵　　　　　　□
464.　'bustling' 熱鬧　　　ŋit⁵⁴ nau⁵⁴
465.　'sharp' (utensil) 快　fai²²

86

466. 'blunt' (utensil) 鈍   tʰin⁵⁴

467. 'fast/quick' 快   fai²²

468. 'slow' 慢   man⁵⁴

469. 'early' 早   tsau³⁵

470. 'late' 晚   an²²   晏

471. 'correct' 對   ŋam⁴²   □

472. 'accurate' 準確   tsʰɔk⁵⁴   □

473. 'wrong' 錯   tsʰɔ²²

474. 'pretty' 漂亮   liaŋ²²   靚

475. 'ugly' 丑   tsʰɐu³⁵   丑

476. 'fat' (animal) 肥   fi³¹

477. 'fat' (human) 胖   fi³¹   肥

478. 'thin' (human) 瘦   sɐu²²

479. 'blind' 瞎   maŋ³¹   盲

480. 'deaf' 聾   luŋ⁴²

481. 'dumb' 啞   ŋa³⁵

482. 'stupid' 傻   muŋ³⁵   懵

483. 'hungry' 餓   ŋɔ⁵⁴   餓

484. 'thirsty' 渴   kiaŋ³⁵ fut⁴²   頸渴

485. 'tired' 累   kʰui⁵⁴   劫

486. 'painful' 疼   tʰuŋ²²   痛

487. 'cheap' (price) 便宜   pʰiaŋ³¹   平

488. 'capable' 有能力   nɐŋ³¹ hai⁵⁴   能□

489. 'very' 很   hau³⁵   好

490. 'more' (comparative) 更   tsuŋ²² ka⁴²   重加

491. 'most' (superlative) 最   tsui²²

492. 'all/both' 都   tɐu⁴²

493. 'only' 只   tsik⁵⁴

494. 'together' 一起   jɐt⁴² tsʰɐi³¹   一齊

495. 'just/precisely' 剛   ŋam⁴²   □

496. 'then' 然後   kɐn⁴² mi²²   跟尾

497. 'then' (emphatic) 就   tsʰɐu⁵⁴

498. 'again' (past) 又   jɐu⁵⁴

499. 'again' (future) 再   tsui²²

500. 'still/yet' 還   tsuŋ²²   重
                     han³¹   還

501. 'also' 也   tɐu⁴²   都

502. 'anyway' 反正   wan³¹ tim⁵⁴   □□

503. 'negation of past actions' 沒有　　mau³⁵　　　　　冇
504. 'not' (general negation) 不　　m̩³¹　　　　　　　唔
505. 'negative imperative' 別　　　m̩³¹ hau³⁵　　　　唔好

4.2.5　Conjunctions, Prepositions, and Particles

506. 'and' 和　　　　　　　　　　　　　　　tʰuŋ³¹　　同
507. 'from' 從　　　　　　　　　　　　　　tsʰuŋ³¹
508. 'progressive/continuous aspect marker'　kin³⁵　　緊
509. 'progressive/continuous aspect marker'　tau²²　　到
510. 'perfective aspect marker' 了　　　　　　hɐu⁵⁴　　後
511. 'disposal/accusative marker' 把　　　　　tsioŋ⁴²　將
512. 'passive marker' 被　　　　　　　　　　pi⁵⁴　　　畀
513. '(by) using' 用　　　　　　　　　　　　kʰai²²　　[扌戒]
514. 'imprecise number' 幾　　　　　　　　　ki³⁵
515. 'possessive marker' 的　　　　　　　　　kɐ²²　　　嘅

4.2.6　Some Remarks

In order to classify major Chinese dialects, Norman (1988: 182) proposes a set of ten key criteria. Out of the ten diagnostic features, eight are related to lexicon. These ten features include:

1) The 3ʳᵈ person pronoun is *tā* 他 or cognate of it.
2) The subordinative particle is *de/di* 的 or cognate to it.
3) The ordinary negative marker is *bù* 不 or cognate to it.
4) The position of the gender marker for animals is prefixed, as in the word for 'hen' *mǔjī*.
5) There is a register distinction only in the *Píng* 平 tonal category.
6) Velars are palatalized before /i/.
7) *Zhàn* 站 or words cognate to it are used for 'to stand'.
8) *Zǒu* 走 or words cognate to it are used for 'to walk'.
9) *Érzi* 兒子 or words cognate to it are used for 'son'.
10) *Fángzi* 房子 or words cognate to it are used for 'house'.

88

To be precise, it is criteria #7 through #10 that are strictly lexical, while criteria #1 through #4 are called by Norman as "grammatical criteria." However, since all of the grammatical criteria more or less involve the choice of cognate, they are also included in discussion and comparison here. Criteria #5 and #6 are strictly phonological and will not be discussed in this section. For any given Chinese dialect, the response to each criterion can be either positive (+), meaning "the statement is true in this dialect", or negative (-), meaning "the statement is not applicable."

Based on the overall responses to all the ten features, Norman classifies Min, Hakka, and Cantonese groups as the Southern group, with negative (-) values for all the criteria. Mandarin dialects in general have positive (+) responses to all the criteria, so they are grouped by Norman as the Northern group. Other dialects have mixed responses to the criteria with positive (+) replies to some criteria and with negative (-) replies to others, and they are called the Centural group, which possessing transitional features between the Northern and the Southern groups.

In the Dapeng dialect, all the criteria related to lexicon have negative values:

1) The 3rd person pronoun is /k^{h}i^{22}/ 佢 (instead of *tā* 他) (§4.2.2).
2) The subordinative (or possessive) particle is /ke^{22}/ 嘅 (instead of *de/di* 的) (§4.2.5).
3) The ordinary negative is /m̩^{31}/ 唔 (instead of *bù* 不) (§4.2.4).

4) The position of the gender markers for animals is suffixed. For male, either /kung⁴²/ 公 or /ku³⁵/ 牯 is used;[44] for female, either /phɔ³¹/ 婆 or /na³⁵/ 嫲 is used. Also, the common diminutive marker /tsɐi³⁵/ 仔, for both human and animals, is postfixed (§4.2.5).
5) (*Phonological feature, irrelevant*.)
6) (*Phonological feature, irrelevant*.)
7) /khi³⁵/倚 (instead of *zhàn* 站) is used for 'to stand' (§4.2.3).
8) /hɐŋ³¹/行 (instead of *zǒu* 走) is used for 'to walk' (§4.2.3).
9) /(a²²) tsɐi³⁵/ 阿仔 (instead of *érzi* 兒子) is used for 'son' (§4.2.1).
10) /uk⁴²/ 屋 (instead of *fángzi* 房子) is used for 'house' (§4.2.1).

According to Norman's (1988) diagnostic test, it is clear that the Dapeng dialect falls into the Southern group together with other major varieties of Chinese, such as Cantonese, Hakka, and Min.

The basic vocabulary list above gives an overall impression that the Dapeng lexicon contains distinctive lexical items from both Cantonese and Hakka, in addition to the large amount of vocabulary shared among Chinese dialects in general. This impression is further confirmed by the results of the picture naming task, which was conducted during the fieldwork.

Based on the lists of distinctive lexical items extracted from Wen (2002, on Hakka) and Zhang (2002, on Cantonese), the picture naming task compiled approximately 100 lexical items that contain different cognates in Hakka and Cantonese. All lexical items were illustrated by pictures, each on one page of an album. Participants

---

[44] According to a senior native speaker of Dapeng, [kung⁴²] 公 and [ku³⁵] 牯 have slightly different meanings. The former refers to male animals in general, sometimes in particular to the uncastrated ones that are raised for breeding purposes, while the latter refers to the juvenile male animals. The distinction between the two female markers, [phɔ³¹] 婆 and [na³⁵] 嫲, is unclear even to that speaker.

were instructed to name the object or action on each page in the Dapeng dialect. The picture naming task was audio recorded, and the original morphemes were identified whenever possible.

After the identification of original morphemes, a parallel study was conducted to compare the cognates used in Dapeng, Hakka, and Cantonese. This study was made through the comparison of original morphemes. Among these approximately 100 distinctive lexical items, the results of comparison show that the Dapeng dialect only shares about 10% of the cognates with Hakka, while more than 80% are shared between Dapeng and Cantonese.

Tables 6 and 7 provide some examples of cognates shared between the Dapeng dialect and either Hakka or Cantonese. Despite the difference in percentile between Hakka-sourced and Cantonese-sourced cognates (roughly 1:8), these tables provide an equal number of examples from both sides for better parallel comparison. Each table contains two nouns, two verbs, two adjectives, one classifier, and one pronoun as examples. Pronunciation and morphemes of Cantonese and Hakka cognates were drawn from one of the following sources: Beijing Daxue (1995, the primary resource), Wen (2002), and Zhang (2002).

	Dapeng		Hakka	
Lexical Item	Morphemes	Pronunciation	Morphemes	Pronunciation
male dog	狗牯	$k\text{ɐ}u^{35}$ ku^{35}	狗牯	$k\varepsilon u^{31}$ ku^{31}
leftover (meal)	菜腳	ts^hui^{22} $ki\text{ɔ}k^{42}$	菜腳	$ts^h\text{ɔ}i^{52}$ $ki\text{ɔ}k^1$
drink (tea)	食	sik^{54}	食	$s\text{ə}k^5$
herd cows	掌牛	$ts\text{ɔ}ŋ^{35}$ $ŋ\text{ɐ}u^{31}$	掌牛	$ts\text{ɔ}ŋ^{35}$ $ŋiu^{11}$
(a pond) dry	潒	lim^{35}	潒	$liam^{35}$
thick (height)	□	$p^h\text{ɐ}n^{22}$	□	p^hun^{44}
CL for people	隻	$tsik^{54}$	隻	$tsak^1$
which	哪	na^{35}	哪	nai^{52}

Table 6. Cognates Shared by Dapeng and Hakka

	Dapeng		Cantonese	
Lexical Item	Morphemes	Pronunciation	Morphemes	Pronunciation
water pit	水凼	sui^{35} $t^h\text{ɐ}m^{31}$	水凼	$ʃ\text{ø}y^{35}$ $t^h\text{ɐ}m^{22}$
Tongue	脷	li^{54}	脷	lei^{22}
slaughter (a pig)	劏	$t^h\text{ɔ}ŋ^{42}$	劏	$t^h\text{ɔ}ŋ^{53}$
chat	傾偈	$k^h\text{ɐ}ŋ^{42}$ $k\text{ɐ}i^{35}$	傾偈	$k^hiŋ^{53}$ $k\text{ɐ}i^{35}$
steep	斜	ts^hia^{31}	斜	$tʃ^h\varepsilon^{33}$
sparse	疏	$s\text{ɔ}^{42}$	疏	$ʃ\text{ɔ}^{53}$
proximal demonstrative	呢	ni^{35}	呢	ni^{55}
CL for quilts	張	$ts\text{ɔ}ŋ^{42}$	張	$tʃ\text{œ}ŋ^{53}$

Table 7. Cognates Shared by Dapeng and Cantonese

In addition, a few lexical items in the Dapeng dialect share cognates with neither Hakka nor Cantonese. For instance, the Dapeng dialect uses the morpheme 剝 /$p\text{ɔ}k^{42}$/ (alternatively, /$m\text{ɔ}k^{42}$/) for the verb "to take off (clothes)", while Hakka uses 脫 /$t^h\text{ɔ}t^1$/ and Cantonese uses 除 /$tʃ^h\text{ø}y^{21}$/. Another example is the verb "to lose (possession)." The

morpheme used in the Dapeng dialect is 落 /lai^{31}/, while both Hakka and Cantonese use

跌, pronounced as /t^{h}iɛt^1/ and /tit^3/, respectively. Overall, Dapeng-specific lexical items

such as these are quite rare, constituting only roughly 5% of the lexical items examined in

the picture-naming task.

In summary, as a Southern Chinese dialect, the Dapeng dialect has a lexicon that,

for the most part, contains distinctive lexical items from both Cantonese and Hakka.

There is considerable difference between the percentiles of Cantonese-sourced versus

Hakka-sourced cognates (roughly 8:1). This difference suggests that the Dapeng lexicon

is much more similar to Cantonese than it is to Hakka.

It should also be noted that the percentiles are based on Mr. L's response,

reflecting the older generation's vocabulary. The contrast between Cantonese-sourced

and Hakka-sourced cognates is even greater. The similarity to Cantonese is even clearer

among the younger generation of Dapeng speakers, especially those in their 20's. While

the younger Dapeng speakers retain most of the Cantonese-sourced cognates in their

vocabulary, some of the Hakka-sourced ones used by the older generation are already

replaced by corresponding cognates from Cantonese or Putonghua. For instance, the verb

"to drink (tea or alcohol)" has the Hakka-sourced cognate 食 /sik^{54}/ in the seniors'

vocabulary. In the younger generation's vocabulary, however, it has been replaced by 飲

/jɐm^{35}/, which is the morpheme used in Cantonese (Beijing Daxue 1995: 373).[45]

---

[45] Although Cantonese is the only Southern Chinese dialect that uses 飲 for "to drink (tea or alcohol)",
it is not an innovation in Cantonese *per se*, but rather a preservation from earlier stages of the Chinese
language, as often seen in classical Chinese texts.

4.3   Syntax

This section summarizes some distinctive syntactic features of the Dapeng dialect. Following the format of the "dialect report", as introduced in §2.3, the primary focus of this section is to discuss the Dapeng syntactic features that are different from those in Standard Chinese, both Putonghua and the Standard Written Chinese. All features were identified from various speaking tasks, both story-telling and spontaneous narrations, and were further supplemented by observations from casual conversations among native speakers of Dapeng during the fieldwork.

Comparisons between the Dapeng dialect and other surrounding Southern Chinese dialects, either Cantonese or Hakka, are the secondary focus of this section. In situations where such comparisons are informative and hence necessary, grammatical rules and example sentences will be cited from different sources for Cantonese (Gao 1980; Matthews and Yip 1994; Yuan 2001; and Zhan 2002) and for Hakka (Hashimoto 1973; He 1993; Yuan 2001; and Li and Xiang 2009).

The description of each feature will be supplemented with example sentences. Each sentence contains orthographic forms (in characters, if identifiable), pronunciation (in IPA), as well as glosses and translation in both Standard Chinese and English, from the top to the bottom.

4.3.1   Word Order: Classifiers

As shown in §4.2.2, the Dapeng dialect has a group of classifiers (CL) while sharing a large portion with Standard Chinese. What makes the Dapeng classifiers distinctive from Standard Chinese in syntax is the "(null) + CL + Noun" structure. That is, the position in front of the classifier in Standard Chinese must be occupied by either numerals or demonstrative adjectives (such as 這 *zhè* and 那 *nà*), unless the classifier immediately follows a verb or a preposition. In the Dapeng dialect, on the other hand, the position can be left empty if the reference to the noun is definite even when the classifier is not at a post-verbal or post-prepositional position.

In contexts where the reference to the noun is clear enough even without the demonstrative adjectives, the Dapeng dialect allows the omission of the numeral or the demonstrative adjective. This feature is shown in Sentence (5), where 隻 /tsik^{54}/ is a "bare classifier" and the reference to the child is specific in the context. In the subsequent sections, Sentences (10), (12), and (23) will provide additional examples of the "(null) + CL + Noun" structure.

(5)   隻   細蚊仔   繼續   去   搵   隻
      tsik^{54}   sɐi^{22} mɐn^{42} tsɐi^{35} kɐi^{22} tsʰɔk^{54} hi^{22} wɐn^{35} tsik^{54}
      CL   child   continue   go   find   CL

      蛤嫲
      kɐp^{42} na^{35}
      frog

      "(這/那)個孩子繼續去找(這/那)隻青蛙"
      'The child continues looking for the frog'

While the omission of demonstratives or numerals is possible in the Dapeng dialect, thus distinguishing it from Standard Chinese, such omission is also allowed in Cantonese (Matthews and Yip 1994, Zhou 1997, Erbaugh 2002, etc.). For instance, in Sentence (5b), the demonstrative in front of the classifier 支 /tʃi⁵⁵/ is omitted, since the reference to the object, pen, is specific and definite in the context.[46] The same phenomenon, in contrast, is not reported for Hakka (Hashimoto 1973, Li and Zhang 1992, Yuan 2001, etc.).

(5b)   Cantonese [47]

支	筆	好	好	寫
tʃi⁵⁵	pɐt⁵	hou³⁵	hou³⁵	ʃɛ³⁵
CL	pen	very	good	write

"(這/那)支筆很好寫"
'The pen is good to write with"

4.3.2   Word Order: Postverbal Adverbs

In Standard Chinese, the common word order of the syntactic construction of verbs and adverbs is "Modifier + Head." In many Southern dialects (such as Cantonese, Hakka, Min, and Wu), in contrast, the "reverse" order of "Head + Modifier" is prevalent (Yue-Hashimoto 1993). For instance, the phrase "先吃飯 *xiān chīfàn* (first-eat-meal, meaning 'to eat meal first')" in Standard Chinese will be normally reordered as "食飯先

---

[46] This dissertation does not aim to discuss in detail the issue of definiteness of bare classifiers in Cantonese and Mandarin. See Cheng and Sybesma (2005), Wu and Bodomo (2009), Tang and Cheng (2014), etc. for further theoretical discussion.

[47] This sentence is cited from Matthews and Yip (1994:93) with Chinese characters added here.

(eat-meal-first)" in both Cantonese and Hakka, with the order of the adverb "first" and the verb phrase "eat-meal" reversed.[48]

The "Head + Modifier" word order is also observed in the Dapeng dialect. Sentences (6) and (7) illustrate two examples below, one with the quantity adverb 多 /to^{42}/ and the other using the scope adverb 埋 /mai^{31}/.

(6) (佢)   又   剝   多   一件   (衫)
     (kʰi^{22}) jɐu^{54} pɔk^{42} tɔ^{42} jɐt^{42} kʰin^{54} (sam^{42})
     3rd SG again   take off   more   one-CL   clothes
     "(他)又再脫了一件(衣服)"
     'He took off one more item of clothing'

(7) (佢)   連   褲   都   剝   埋
     (kʰi^{22}) lin^{31} fu^{22} tu^{42} pɔk^{42} mai^{31}
     3rd SG even   pants   also   take off   along/in addition
     "(他)連褲子都一併脫了"
     'He even also took off his pants'

Despite the by-default, "reversed" word order, the Dapeng dialect also allows the "non-reversed" order, which places the adverbial modifier before the verbal head. In the narration of the story of *The North Wind and the Sun*, for instance, the temporal adverb 先 /sin^{42}/ appears in both the preverbal and the postverbal positions. Sentence (8) shows an example of the preverbal 先 /sin^{42}/. It should be pointed out that the distribution of the preverbal and the postverbal 先 /sin^{42}/ does not show any clear correlation with age as some other morphosyntactic features do: both usages occur across generations.

---

[48] In Cantonese and Hakka, the pronunciation of this phrase is /ʃik^2 fan^{22} ʃin^{53}/ and /sət^5 fan^{52} sian^{44}/, respectively.

(8)　北風　　　　　就　　　　　　講　睇　我　先　來
　　pɐk⁴² fuŋ⁴²　tsʰɐu⁵⁴　　　kɔŋ³⁵ tʰɐi³⁵ ŋɔ³⁵ sin⁴² lui³¹
　　north-wind　　then　　　　say　look　1ˢᵗ SG　first　come
　　"北風就說'看我先來'"
　　'The north wind then says, 'look at me, let me try first''

Sometimes the preverbal and the postverbal 先 /sin⁴²/ can even co-occur in one sentence. Sentence (9) was uttered by a 40-year old female speaker. It demonstrates the relatively flexible position of the verbal 先 /sin⁴²/.

(9)　北風　　　　　帶頭　　　　先　來　比試　　　先
　　pɐk⁴² fuŋ⁴²　tai²² tʰɐu³¹　sin⁴² lui³¹ pi³⁵ si²²　　sin⁴²
　　north-wind　　take-the-lead　first　come　try　　　first
　　"北風先帶頭來嘗試"
　　'The north wind took the lead and tried first'

4.3.3　Word Order: Disposal Construction

The disposal construction in Chinese expresses a meaning of "affectedness." In this grammatical construction, the action or influence acted upon a nominal is emphasized as the nominal (the "affectee") is moved before the verb. The Dapeng dialect uses the preposition 將 /tsiɔŋ²²/ to introduce the affectee, which derives from a verb meaning "to take." The basic structure is "S + 將 TSIONG + O + V", which indicates "the object is affected or disposed by the subject in the manner of the action V."

(10)　條　狗　將　　　　竇　蜂　整　落來
　　　tʰiu³¹ kɐu³⁵ tsiɔŋ²²　tɐu²² fuŋ⁴² tsɐŋ³⁵ lɔk⁵⁴ lui³¹
　　　CL　dog　TSIONG　　CL　wasp　do　fall-come
　　　"（那）條狗把（那）窩蜜蜂弄下來"
　　　'(That) dog takes down (that) swarm of wasp'

(11) 農民　　　將　　　　　一籮　　　　雪梨　　　　運
nɔŋ³¹ mɐn³¹　tsiɔŋ²²　　jɐt⁴² lɔ³¹　　sit⁴² li³¹　　wɐn⁵⁴
peasant　　TSIONG　　one-CL　　　pear　　　　carry

返　　屋企
fan⁴²　uk⁴² kʰi²²
back　home

"農民把一筐雪梨運回家"
'The peasant carries back home a basket of pear'

The use of 將 /tsiɔŋ²²/ in the Dapeng dialect is largely equivalent to the use of the same morpheme 將/tʃœŋ⁵⁵/ in both Cantonese and Hakka. The use of 將/tʃœŋ⁵⁵/ as a disposal marker is similar to, and yet more restricted than, the well-known use of 把 *bǎ* in Putonghua. In Cantonese 將/tʃœŋ⁵⁵/ is usually used in cases of movement or removal of objects, and it is applicable in both physical and metaphorical cases (Matthews and Yip 1994: 144). Hakka also uses 將 /tsiɔŋ⁴⁴/, and according to Yuan (2001: 173), it is used in a similar way as in Cantonese. Sentences (11b) and (11c) show examples of disposal construction in Cantonese and Hakka from these two studies, respectively.

(11b) Cantonese

佢　　將　　　　　□　　　　污糟　　　　衫　　周圍　　　　□
kʰøy³⁵ tʃœŋ⁵⁵　　ti⁵⁵　　u⁵³ tʃou⁵³　ʃam⁵³ tʃɐu⁵³ wɐi²¹　pʰɛk²
3ʳᵈ SG TSIONG　CL　　dirty　　　clothes around　　　throw

"他/她把(那)些臟衣服到處扔"
'He/she throws his dirty clothes all over the place"

(11c)   Hakka

佢	將	茶杯	打	爛	□
ki^{11}	tsiɔŋ^{44}	tsʰa^{11} pi^{44}	ta^{31}	lan^{52}	e^{31}
3rd SG	TSIONG	tea-cup	break	broken	PFV

"他/她把茶杯打破了"
'He/she broke the tea cup"

In addition to the basic structure "S + 將 TSIONG + O + V", both Cantonese and Hakka also have a special structure of disposal construction: "S + 將 TSIONG + O$_i$ + V + Pronoun$_i$", in which case the pronoun and the object have the same reference (Chappell 2007: 10). Sentence (11d) is an example in Cantonese from Chappell's study. The Dapeng dialect, based on field data from the speaking tasks, does not show any case of using this structure. None of the native speakers, regardless of their age, used this structure. It seems possible that this structure is not allowed in the Dapeng dialect. If so, this structure distinguishes the Dapeng dialect from both Cantonese and Hakka.[49]

(11d)   Cantonese

將	□	頭髮	染	黑	佢
tʃœŋ^{55}	ti^{55}	tʰɐu^{21} fat^{33}	im^{23}	hak^{5}	kʰøy^{35}
TSIONG	CL	hair	dye	black	3rd SG

"把那些頭髮染黑"
'Dye that/this hair black"

---

[49] It is also possible that the Dapeng dialect does have this structure, but is simply not reflected in the database of the current study. This could be a focus for future fieldwork and a topic for future studies.

There also appears to be an age-related usage of 將 /tsiɔŋ²²/ in the Dapeng dialect. Speakers younger than 40 tend to use this construction notably more frequently than those 40 and older.

4.3.4   Word Order: Passive Construction

The passive construction in the Chinese language places a preposition in front of the agent of an action and makes the recipient the subject of the sentence. Standard Chinese mainly uses the preposition 被 *bèi*, which originally means "to suffer (from)", to introduce the agent.[50] Unlike Standard Chinese, the Dapeng dialect uses another preposition 俾 /pi³⁵/, which is also the most common passive marker in Cantonese (although pronounced differently as /pei³⁵/). In Hakka, the preposition 分 /pun⁴⁴/ is normally used.

The different use of passive markers has a typological significance in the overall context of languages in China. Chappell (2015b: 27-36) proposes seven diachronic sources of passive markers for Sinitic languages, which point back to the original, verbal meaning of the modern passive markers.

The passive markers in Mandarin belong to several sources: Type I ("suffer", in the case of 被 *bèi*), Type V ("give", in the case of 給 *gěi*), and Type VI ("speech act verbs "tell", "call", "ask", in the cases of 叫 *jiào* and 讓 *ràng*). In contrast, the passive

---

[50] There are several other passive markers in Mandarin, including 叫 *jiào*, 讓 *ràng*, and 給 *gěi*. According to Li and Chen (2005: 3), the frequency of occurrence of these passive markers is very different, ranking as 被 *bèi* > 叫 *jiào*/讓 *ràng* >> 給 *gěi*. In this section, 被 *bèi* is viewed as the representative of the Mandarin passive markers and is compared with the Dapeng passive marker 俾 /pi³⁵/.

markers in Cantonese and Hakka belong to just one source: Type V ("give", in the cases of 俾 /pei³⁵/ in Cantonese and 分 /pun⁴⁴/ in Hakka). Apparently 俾 /pi³⁵/ in the Dapeng dialect should also fall into Type V.

Despite the typological differences in diachronic sources, the action indicated by the passive construction is normally an undesirable one among Mandarin, Cantonese, Hakka, and Dapeng, regardless of the choice of passive markers. Sentence (12) provides an example, where "being bitten by the nesokia" involves an unpleasant and undesirable situation.

(12) 細蚊仔   俾 隻 地鼠   咬 到 隻
   sɐi²² mɐn⁴² tsɐi³⁵ pi³⁵ tsik⁵⁴ tʰi⁵⁴ si³⁵ ŋau²² tau²² tsik⁵⁴
   child   PASS[51]CL nesokia bite RES[52] CL

   面
   min⁵⁴
   face

   "孩子被地鼠咬到臉"
   'The child is bitten by the nesokia on the face'

A major significant distinction between the 俾 /pi³⁵/ construction in the Dapeng dialect and the 被 *bèi* construction in Standard Chinese is the compulsory occurrence of the agent. In Standard Chinese, the agent of an action is not as crucial and may be excluded from the passive construction;[53] in the Dapeng dialect, however, similar to

---

[51] "PASS" refers to passive markers.
[52] "RES" refers to resultative verbs.
[53] To be precise, the inclusion or exclusion of the agent depends on the choice of passive marker in Standard Chinese. According to Li and Chen (2005: 2), both 被 *bèi* and 給 *gěi* constructions allow the exclusion of the agent. The difference is in genra: in the 被 *bèi* construction, the exclusion of the agent most likely happens in narration forms; in the case of the 給 *gěi* construction, the exclusion of the agent

Cantonese (Matthews and Yip 1994: 149) and Hakka (Huang 2015: 170), the agent has to appear. That is, the main verb has to be preceded by the object of the passive construction, as shown in Sentence (13).

(13) 你　　係唔係　　　　　俾　人　笑
　　　ni^{22}　hɐi^{54} m̩^{31} hɐi^{54}　　pi^{35}　jɐn^{31}　siu^{22}
　　　2nd SG be-NEG^{54}-be　　PASS　people　laugh (at)
　　　"你是不是被人笑話"
　　　'Is it that you are laughed at?'

According to Dapeng speakers' judgment, as in Cantonese, it is illegal to omit the agent 人 /jɐn^{31}/ ("people") in this situation. Although the real agent of the action "to laugh at" is not necessarily clear in this situation, a non-specific term "people" is still required.

4.3.5　Word Order: Comparative Construction

In situations where two objects are different under comparison, the Dapeng dialect primarily uses the comparative markers 過 /ko^{22}/. The original, literal meaning of this word is "to pass." In the comparative construction it is grammaticalized and functions as a preposition. In the Dapeng dialect, the word order is "X + adjectival predicate + 過 ko^{22} + Y", which is also seen in Cantonese (Matthews and Yip 1994: 166) and Hakka (He 1993: 72). Standard Chinese uses another comparative marker, 比 *bǐ*, and

---

could happen in both narration and dialog forms. The 叫 *jiào*/讓 *ràng* construction does not allow such exclusion in either situation.
[54] "NEG" refers to negation.

the corresponding structure is "X + 比 *bǐ* + Y + adjectival predicate." Interestingly, Hakka also allows a somewhat mixed structure with both 過 /kuɔ^{52}/ and 比/pi^{31}/, namely, "X + 比/pi^{31}/ + Y + 過 /kuɔ^{52}/ + adjectival predicate" (Yuan 2001: 173).

According to Chappell's (2015b: 37-45) proposal of the seven structural types of comparatives in Sinitic languages, Standard Chinese falls into Type I, "Prepositional Comparative", a.k.a. the "Mandarin Type" for its prevalence in Northern China. The structural configuration of Type I is "NP$_i$ [CM NP$_j$] VP",[55] with the comparative marker being a part of the prepositional phrase combined with NP$_j$.

Cantonese, on the other hand, belongs to Type II, "Transitive Comparative", which is widespread among dialects in Southern and Southwestern China. The structural configuration of Type II is "NP$_i$ VP CM NP$_j$", the comparative marker acting as the complement of VP.

While Hakka also belongs to Type II, it also fits in Type V, "Hybridized Comparative", which is a combination of Type I and Type IV ("Adverbial Comparative"). The structural configuration of Type V is "NP$_i$ [CM$_a$ NP$_j$] CM$_b$ VP", with the first half adapted from Type I ("NP$_i$ [CM NP$_j$]"). In the second half, the comparative marker (CM$_b$) is essentially an adverb with the meaning of "more." Such hybridity could be a result of interdialect contact (Chappell's 2015b: 41, also containing examples).

Apparently 俾 /pi^{35}/ in the Dapeng dialect should fall into Type II together with Cantonese. Sentences (14) and (15) shows two examples.

---

[55] "CM" refers to comparative marker.

(14)　　我　　　高　　　過　　　　　　佢
　　　　ŋɔ^{35}　kau^{42}　kɔ^{22}　　　　kʰi^{22}
　　　　1st SG　tall　　COMP　　　　3rd SG
　　　　"我比他高"
　　　　'I am taller than him'

(15)　　我　　　本領　　　　　大　　　過　　　你
　　　　ŋɔ^{35}　pun^{35} liaŋ^{22}　tʰai^{54}　kɔ^{22}　ni^{22}
　　　　1st SG　ability　　　　big　　COMP　2nd SG
　　　　"我的本領比你大"
　　　　'I am more capable than you'

Both the picture naming and film narration speaking tasks, however, also show some variation across age. Some native speakers from the younger generation (roughly below the age of 35) also use another comparative markers, 比 /pi^{35}/, which is identical to the comparative construction in Standard Chinese. In correspondence, the word order becomes "X + 比 pi^{35} + Y + adjectival predicate", as illustrated in Examples (16) and (17).

(16)　　我　　　比　　　　　　佢　　　高
　　　　ŋɔ^{35}　pi^{35}　　　　kʰi^{22}　kau^{42}
　　　　1st SG　COMP　　　　3rd SG　tall
　　　　"我比他高"
　　　　'I am taller than him'

(17)　我　　本領　　　比　　你　　大
　　　ŋɔ³⁵　pun³⁵ liaŋ²²　pi³⁵　ni²²　tʰai⁵⁴
　　　1ˢᵗ SG ability　　　COMP 2ⁿᵈ SGbig
　　　"我的本領比你大"
　　　'I am more capable than you'

In situations where two objects are equal in quality or quantity, the Dapeng dialect uses the same structure as the majority of the Chinese varieties: "X + conjunction + Y + adverb + adjectival predicate." The main difference lies in the choice of the adverb, which indicates "equality." In Standard Chinese, the adverb is 一樣 *yíyàng*; in Dapeng, its counterpart is 平 /pʰeng³¹/, literally meaning "being level, even." This is shown in Sentence (18)

(18)　我　　同　　佢　　平　　高
　　　ŋɔ³⁵　tʰuŋ³¹　kʰi²²　pʰɐŋ³¹　kau⁴²
　　　1ˢᵗ SG and　　3ʳᵈ SG same　tall
　　　"我和他一樣高"
　　　'I am as tall as him'

In some cases, the object of comparison can be omitted if the context has made this information clear. The structure "X + 比較 *bǐjiào* + adjectival predicate" is shared by the Dapeng dialect, Standard Chinese, and the other two major Chinese dialects: Cantonese and Hakka. Sentence (19) is an example from the story telling task.[56]

---

[56] There should be some other strategies to make comparison in the Dapeng dialect. Due to the limited amount of data, the current study does not cover a few forms of comparison, for instance:
　(a)　Degrees of comparison, as in English "my car is *twice* bigger than yours";
　(b)　Negative comparison, as in English "he is not as tall as his father."
　These strategies ought to be discussed in future studies when more fieldwork data are available.

(19) 哪人　　　　　比較　　　　能□
na^{54} jɐn^{31}　　pi^{35} kau^{42}　　nɐŋ^{31} hai^{54}
who　　　　　compare　　　capable
"誰比較厲害"
'Who is more capable?'

4.3.6　Aspect: Perfective

As in many other varieties of Chinese, 後 /heu^{54}/ in the Dapeng dialect first means "back" (noun) and derives from it the postpositional function "after." What makes 後 /heu^{54}/ in the Dapeng dialect different from both Hakka and Cantonese and from many other dialects, however, is that it has also been fully grammaticalized into a perfective (PFV) aspect marker, as demonstrated in Sentences (20) and (21).

(20)　(農民)　　　　摘　　生果　　　　摘後　　　　兩籮
(nɔŋ^{31} mɐn^{31})　tsak^{42}　saŋ^{42} kɔ^{35}　tsak^{42} hɐu^{54}　liɔŋ^{35} lɔ^{31}
(peasant)　　pick　fruit　　　pick-PFV　　two-CL
"(農民)摘水果摘了兩籮筐"
'(The peasant) picked two baskets of fruit'

(21)　細蚊仔　　　　　俾後　　　　幾隻　　　　生果　　　　俾
sɐi^{22} mɐn^{42} tsɐi^{35}　pi^{35} hɐu^{54}　ki^{35} tsik^{54}　saŋ^{42} kɔ^{35}　pi^{35}
child　　　　　give-PFV　　several-CL　fruit　　　　give

佢哋　　　　食
kʰi^{22} tʰi^{54}　sik^{54}
3rd pl　　　eat

"孩子給了幾個水果給他們吃"
'The child gave them some fruit to eat'

107

In additional to 後 /heu⁵⁴/, the Dapeng dialect also has another perfective marker, 咗 /tsɔ³⁵/, which is also the perfective marker in Cantonese.[57] While 後 /heu⁵⁴/ is used extensively in the Dapeng dialect across all generations, the use of 咗 /tsɔ³⁵/ is in general restricted to the young generation (below the age of 30). There is also occasional usage of 咗 /tsɔ³⁵/ observed among the oldest group of speakers (above the age of 65), but the amount is minimal.

(22)　　(子女)　　　　過咗年　　　　　　返嚟
　　　　(tsi³⁵ ni³⁵)　　kɔ²² tsɔ³⁵ nin³¹　　fan⁴² lɐi³¹
　　　　(children)　　pass-PFV-year　　return-come
　　　　"(子女)過了年回來"
　　　　'The children return (home) after the New Year'

4.3.7　Aspect: Imperfective (Progressive vs. Continuous)

The Chinese language in general makes a distinction between progressive (PROG, indicating dynamic meanings) and continuous (CONT, indicating static meanings) aspects. For instance, Mandarin distinguishes 在 *zài* (PROG) from 著 *zhe* (CONT) (Klein et al. 2000), Cantonese distinguishes 緊/kɐn³⁵/ (PROG) from 住/tʃy²²/ (CONT) (Zhan 2002), and Hakka distinguishes 緊/kin²⁴/ (PROG) from 等/ten²⁴/ (CONT) (Song 2008).

In the Dapeng dialect, however, these two imperfective aspects are not distinguished. The same aspect marker, 緊 /kɐn³⁵/, is used to denote both dynamic actions

---

[57] Its corresponding perfective marker in Hakka is □ /e³¹/ (Hashimoto 1973: 443).

and static states. Sentences (23) and (24) demonstrate how the post-verbal marker 緊 /kɐn³⁵/ indicates both the duration of the state "being put (somewhere)" and that of the action "chasing."

(23)　(佢)　養到　　　　隻　　蛤嬤 …　　　就　　　　　放緊
　　　 (kʰi²²) iɔŋ³⁵ tau²²　 tsik⁵⁴　kɐp⁴²na³⁵…　tsʰɐu⁵⁴　　 fɔŋ²² kɐn³⁵
　　　 (3ʳᵈ SG) raise-CONT　CL　　frog　　　　then　　　　　put-CONT

　　　在　　房間
　　　tsʰui⁵⁴　fɔŋ³¹ kan⁴²
　　　in　　　room

　　　"(他)養著一隻青蛙 … 就(把它)放在房間裡"
　　　'(He) has a frog… and puts it in the room'

(24)　貓頭鷹　　　　　　　追緊　　　　佢哋
　　　miu³⁵ tʰɐu³¹ jɐŋ⁴²　　tsui⁴² kɐn³⁵　kʰi²² tʰi⁵⁴
　　　Owl　　　　　　　　chase-PROG　3ʳᵈ PL
　　　"貓頭鷹在追他們"
　　　'The owl is chasing them'

It should also be noted that Example (23) showed two continuous aspect markers. In addition to 緊 /kɐn³⁵/, another marker 到 /tau²²/ is also occasionally used. Sentences (25) and (26) below show more cases where 到 /tau²²/ marks the continuous aspect. Sentence (26) is of particular interest, in which 緊 /kɐn³⁵/ and 到 /tau²²/ are used interchangeably in two nearly identical, repetitive clauses.

(25) (嗰□)　　　□到　　　一隻　　　鹿
 kɔ³⁵ liaŋ⁵⁴　　piaŋ²² tau²²　jɐt⁴² tsik⁵⁴　　luk⁵⁴
(that-place)　　hide-CONT　one-CL　　　deer
"(那兒)藏著一隻鹿"
'(There) hides a deer'

(26) (佢)　牽緊　　　　一隻　　　　羊仔，　　　牽到
(kʰi²²) hin⁴² kɐn³⁵　jɐt⁴² tsik⁵⁴　iɔŋ³¹ tsɐi³⁵　hin⁴² tau²²
(3ʳᵈ SG) lead-CONT　one-CL　　　goat-little　　lead-CONT

隻　　　羊仔　　　　在　　□　　　　　經過
tsik⁵⁴　iɔŋ³¹ tsɐi³⁵　tsʰui⁵⁴ liaŋ⁵⁴　　kɐŋ⁴² kɔ²²
CL　　goat-little　　at　　(that) place　pass by

"(他)牽著一隻小羊，牽著一隻小羊在那兒經過"
'(He) passes by that place leading a little goat with a rope'

4.3.8　Some Remarks

It has been shown above that the Dapeng dialect has a unique syntactic feature, i.e. the merging of the progressive and continuous imperfective aspects. The other features are still highly distinctive from Standard Chinese, but they all resemble either Cantonese or Hakka in various degrees.

Between these two neighboring Southern dialects, the syntactic resemblance between the Dapeng dialect and Cantonese is slightly more significant, especially so in word order. There are also variations across generations, with the younger generation adopting features both from Cantonese (e.g. perfective aspect marker) and from Standard Chinese, such as comparative construction.

4.4   Summary

Compared with the sound system, the Dapeng lexicon and syntax show more similarity to both Cantonese and Hakka. This confirms the general observation in Chinese dialectology that the differences among Chinese dialects are most significant in phonology, then in lexicon, and the least in syntax.[58] On the other hand, there are more intergenerational variations in the Dapeng lexicon and syntax than in the sound system.

Overall, the Dapeng lexicon and syntax show more similarity to Cantonese. In terms of intergenerational variations, the general direction of change also leans towards Cantonese (and sometimes towards Standard Chinese in syntax) as age decreases.

---

[58] This, again, is the reason why the "dialect report" focuses more on phonology.

Chapter 5   Koineization: The Formation Process

Based on the language facts presented in the previous chapters (the results of

Hakka-Cantonese contact), this chapter proceeds to analyze the processes (or

mechanisms) of the contact, which contributed to the formation of the current-day

Dapeng dialect. In this chapter, Trudgill's (1986) model of "koineization" will be

discussed and then applied to account for the processes that gave rise to the Dapeng

dialect. Both linguistic and socio-historical features that were described or introduced in

previous chapters will be re-examined in this chapter and will serve as evidence to

support the proposed process of "koineization."

First, a literature review in §5.1 introduces some key concepts. Those concepts

that have caused confusion in terminology are revisited and redefined. §5.2 shows that

Siegel's (1985) category of "immigrant koine" is particularly relevant to the case of

Dapeng and emphasizes the importance of considering the issue of migration in the study

of Chinese dialects. After that, the immigration history of Dapeng is examined in §5.3,

which provides socio-historical evidence for the hypothesized koineization process of the

Dapeng dialect. Then, it will be further supported by linguistic evidence. §5.4 further

supports the hypothesis from a linguistic perspective by providing evidence from the

structure of Dapeng.

5.1   Koineization: A Literature Review

In the literature, the terms pertaining to koineization have not been clearly defined. This section thus starts with a brief review of some of the key terms, in order to analyze the Dapeng data and hypothesize the process of Dapeng formation with less ambiguous, better clarified terminology.

5.1.1   The Concept of "Koine" (Siegel 1985)

According to Siegel (1985: 358), the term **"koine"** refers to the result of contact between dialects such as regional dialects. "Koine" originates from the Greek word for "*common*", and was originally used as the name of a particular variety of ancient Greek, which was a consequence of the mixing of several Greek regional dialects and was once the *lingua franca* of the eastern Mediterranean. Recently the use of this term has been extended from the Greek context to generally referring to a type of language variety that is "the stabilized result of mixing of linguistic subsystems such as regional or literary dialects" (Siegel 1985: 363). In terms of structure, a koine often comprises linguistic features from different source dialects; in terms of function, it usually serves as a *lingua franca* among speakers of those dialects.

He also distinguishes two types of koines: regional koine and immigrant koine. The **regional koine** refers to the outcome of the contact of two geographically adjacent dialects, which usually does not replace the function of either dialects (that is, the koine remains as no more than a tool for inter-dialectal communication). The original Greek koine is an example of a regional koine.

113

The **immigrant koine**, on the other hand, emerges in new settlements established by migrants. Unlike the regional koine, the immigrant koine often replaces the original contributing dialects and becomes the primary language of the newly established community (Siegel 1985: 363-364). An example of immigrant koine is the new variety of English in the Town of Milton Keynes, where the majority of the population originally came from Greater London and other parts of the United Kingdom. After the new town was officially designated in 1967 (Kerswill and Williams 2000: 78). Unless otherwise specified, all issues regarding koine formation discussed below fall into the second category, the immigrant koine.

According to Siegel (1985: 358; cf. Kerswill 2013: 520), the koine is characterized by **reduction** and **simplification**. Although both are referred to as processes in Siegel (1985), in a later study he refines his terminology. Following Trudgill (1986), Siegel (2001: 176) keeps simplification as a process and refers to the other one as levelling, which is associated with the attrition of input dialect variants. That is, he agrees that the process he described earlier as "reduction" is essentially levelling, a more general linguistic process, and that the term "reduction" should be viewed only as the result of levelling. In the rest of this chapter, the term "reduction" is used consistently as the result of the processes that lead to the elimination of competing features (levelling) or irregularities (simplification) in the koine.

5.1.2   The Original Definition of Koineization (Trudgill 1986)

To study the process of koine formation, the model "**koineization**" was developed by a group of scholars (Siegel 1985, 2001; Trudgill 1986; Britain and Trudgill 1999; Trudgill et al. 2000; Kerswill and Williams 2000; among many others). Koineization results in koines, which shows "structural convergence between closely related linguistic systems, eventually leading to the stabilization of some compromise variety" (Hinskens 2001: 200; cf. Hinskens, Auer, and Kerswill 2005:11). In terms of processes that koineization involves, most of the recent studies have followed Trudgill's (1986) model. In this classic, widely-cited study of dialect contact, Trudgill (1986: 107-108) summarizes the role of koineization in new-dialect formation as follows.

> In dialect contact and dialect mixture situations there may be an enormous amount of linguistic variability in the early stages. However, as time passes, focusing takes place by means of a reduction of the forms available. This *reduction* takes place through the process of koineization, which consists of the *levelling* out of minority and otherwise marked speech forms, and of *simplification*, which involves, crucially, a reduction in irregularities. (The degree of simplification, and possibly its nature, may be influenced by lingua franca usage (pidginization) and by language death in situations which involve language contact as well as dialect contact.) The result of the focusing associated with koineization is a historically mixed but synchronically stable dialect which contains elements from the different dialects that went into the mixture, as well as interdialect forms that were present in none (italics in original)

Apparently Trudgill does not treat "koineization" as a unique process of linguistic change which is independent of other processes; rather, it consists of the more general, independent linguistic processes of **levelling** and **simplification**. He defines the process

115

of levelling as involving the loss of marked and/or minority variants, which are less normal and more irregular, and the process of simplification as the reduction of phonological and morphophonemic complexity and irregularities (Trudgill 1986: 126, Kerswill 2010: 231). Both processes lead to a reduction (which, again, suggests a result) of the forms available in the contributing dialects that are in contact. As the result of levelling, reduction indicates that some competing variants in the input dialects are eliminated in favor of others; as the result of simplification, reduction indicates the elimination of structural complexity and irregularities.

A fundamental distinction exists between the two processes. Levelling contributes to the reduction of intersystemic variations (in Hinskens, Auer, and Kerswill's terms, 2005: 2), i.e. those different linguistic variations *between* contributing dialects. Simplification leads to the reduction of intrasystemic variations, i.e. the structural complexity and irregularity *within* the emerging koine. In other words, the locus of simplification is in the rudimentarily converged dialect, which results from the previous process of levelling. In this sense, the outcome of levelling is very similar to dialect convergence as discussed by Hinskens, Auer, and Kerswill (2005), both preparing ways for further simplification.

Siegel (1985: 370) attempts to distinguish koineization from dialect convergence by arguing that the former "leads to the development of a new compromise variety with features of the contributing varieties," whereas the latter "leads to changes in the contributing varieties themselves without development of a new variety." However, the difference between these two concepts could have been exaggerated in his terminology.

116

While focusing on the *results* of the processes of koineization versus dialect convergence, Siegel overlooks the similarities in the *processes* involved. On the other hand, Hinskens, Auer, and Kerswill (2005) see the similarity between koineization and dialect convergence in that both are outcomes of general linguistic processes in which structural differences between/among the source dialects are levelled out.

In fact, there are two types of dialect convergence: one dialect converging to another (and losing its own features) and two dialects converging to each other (and resulting in a compromised dialect). Siegel (1985) assigns "dialect convergence" narrowly to the first type and "koineization" to the second. By contrast, Hinskens, Auer, and Kerswill (2005) treat "dialect convergence" as a cover term for both types. Given favorable socio-historical environments, two dialects can potentially converge towards each other, forming a new product (a koine) that may eventually replace the functions of the original contributing dialects. That also explains why they call the koine "the results *par excellence* of dialect convergence" (Page 12). The two different ways of defining "dialect convergence" vis-à-vis "koineization" is compared in Table 8. (A and B are both contributing dialects that are in contact.)

		Siegel (1985)	Hinskens et al. (2005)	
Type 1	A → B	Dialect convergence		
Type 2	A B } → C	Koineization	Dialect convergence	Koineization

Table 8. Different Definitions of Dialect Convergence and Koinezation

117

5.1.3 Processes of New Dialect Formation

In Trudgill's (1986) model, koineization is actually one of the stages of a larger, more complicated phenomenon, i.e. **new dialect formation**. In the fuller model, Trudgill also discusses what happens both before and after levelling and simplification, the two processes associated with koineization. For clarity, below is a full citation of Trudgill's (1986: 126) summary of the processes and the results of new dialect formation.

> We can now summarize our findings as follows. In a dialect mixture situation, large numbers of variants will abound, and, through the process of *accommodation* in face-to-face interaction, *interdialect* phenomena will begin to occur. As time passes and *focusing* begins to take place, particularly as the new town, colony, or whatever begins to acquire an independent identity, the variants present in the mixture begin to be subject to *reduction*. Again this presumably occurs via accommodation, especially of salient forms. This does not take place in a haphazard manner, however. In determining who accommodates to whom, and which forms are therefore lost, demographic factors involving proportions of different dialect speakers present will clearly be vital. More importantly, though, more purely linguistic forces are also at work. The reduction of variants that accompanies focusing, in the course of *new-dialect formation*, takes place via the process of *koineization*. This comprises the process of *levelling*, which involves the loss of marked and/or minority variants; and the process of *simplification*, by means of which even minority forms may be the ones to survive if they are linguistically simpler, in the technical sense, and through which even forms and distinctions present in all the contributory dialects may be lost. Even after koineization, however, some variants left over from the original mixture may survive. Where this occurs, *reallocation* may occur, such that variants originally from different regional dialects may in the new dialect become *social-class dialect variants, stylistic variants, areal variants*, or, in the case of phonology, *allophonic variants* (italics in original).

In light of Trudgill's summary, we can now describe the procedures of new dialect formation and discuss the relationship between koineization and the multi-stage new dialect formation. First, in order to accommodate each other, speakers of different dialects, either consciously or unconsciously, identify the differences between two or more contributing dialects in the mixture situation. As they also try to accommodate to each other's speech, some form of interdialect begins to occur. This is the first stage, that of accommodation.

The second stage of new dialect formation is koineization, which includes the linguistic processes of levelling and simplification. The process of levelling begins as speakers of different dialects try to eliminate these distinctions by keeping the similarities between these two linguistic systems and create a shared system. Then this newly emerged system serves as the preliminary and yet critical medium of communication, while speakers of the different dialects continue to further reduce the phonological and morphophonemic complexity of the emerging koine. This process is called simplification, another linguistic process through which the new linguistic system becomes stabilized.

However, as Trudgill suggests, massive levelling and simplification of this type will not occur until face-to-face interaction among speakers of different source dialects in a stabilized co-inhabitance accumulates to a certain degree. The critical point is when a new, independent identity begins to be acquired by all groups in the new settlement, for instance a new town or a colony. The new identity motivates a "reduction in the number of variant forms and the increase in sociolinguistically predictable variation, that is, the

119

(re-)emergence of norms [and stability]" (Kerswill 2010: 230; Kerswill and Trudgill 2005: 199).

After the initial formation of the new identity, **focusing** begins to take place. Focusing describes a situation when members of a language community feel their own language to be clearly distinct from other languages by some delineated boundaries in between and when they have a high level of agreement on what belongs to the language and what does not. A common identity shared within a speech community is the primary factor that contributes to focusing, especially in new settlements that are distant or isolated from the source communities (Trudgill 1986: 85-86, c.f. Tuten 2003: 39-41). Focusing, therefore, is clearly not a linguistic process but a sociolinguistic one.

The sociolinguistic process of focusing also sheds light on the distinction between koineization and dialect levelling in general. While levelling is one of the two key linguistic processes involved in koineization, focusing emphasizes the importance of a specific socio-historical circumstance for koineization to happen, viz. a newly formed, independent, unified identity among speakers from different dialect groups (for instance, in the case of Fiji as reported by Siegel 1985). That is to say, koineization is a special result of the combination of the two common linguistic processes—dialect levelling and simplification—accompanied by focusing, which is a sociolinguistic process triggered by specific socio-historical factors. Following Hinskens, Auer, Kerswill's (2005) logic, one could say that koineization is the result *par excellence* of dialect levelling given specific socio-historical environments.

Through the processes of levelling and simplification, a possible outcome of koineization, named **reallocation** may follow. This is the fourth step, again not a linguistic process, but rather a sociolinguistic one. Britain and Trudgill (1999: 245) define reallocation as a process in which "two or more variants in the dialect mix survive the levelling process but are refunctionalised, evolving new social or linguistic functions in the new dialect." In other words, while some variants win out in levelling, some others left over from the original pool of the dialect mixture can still be available to the new dialect to be repurposed in new ways.

Where this happens, reallocation may occur, and such left over variants become social-class variants or stylistic variants in the emerging dialect (Kerswill and Trudgill 2005: 199). Accordingly, it seems fair to define reallocation as both a sociolinguistic process and as the result that it causes.[59] Since the issue of reallocation is not relevant to the current study of Dapeng, which discusses very little regarding sociolinguistic variations in this dialect, this social process and its consequences will not be further addressed.

Trudgill's (1986) model may be better understood by the illustration in Figure 5. As indicated by the shaded area in this figure, koineization is the second stage of new dialect formation.

---

[59] For more details concerning reallocation, see Britain and Trudgill (1999).
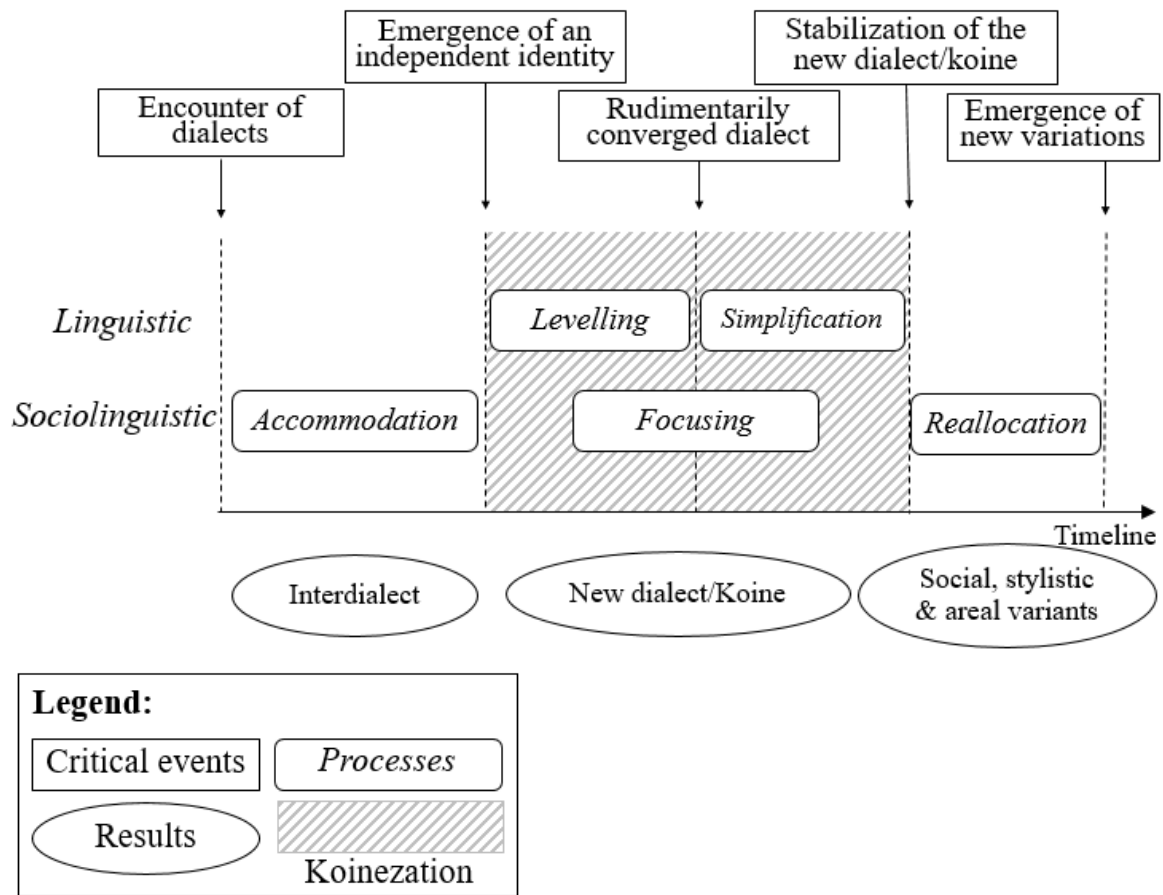
Figure 5. A Graphic Illustration of Trudgill's (1986) Model of New Dialect Formation

Figure 5 places special emphasis on the distinction between processes and results, which seems somewhat vague in Trudgill's terminology. It also distinguishes linguistic processes from sociolinguistic ones. As dialects encounter one another in a new settlement, speakers start to accommodate to each other in their speech, and such accommodation leads to an unstable, prekoine interdialect. As time progresses, a local, independent identity that is closely tied to the new community emerges, and it begins to be acquired by and shared among all dialect speakers from different groups. This is when koineization starts to take place.

As Figure 5 illustrates, koineization *per se* is not a unique linguistic process but consists of two more basic linguistic processes, namely, levelling and simplification. Both processes are triggered by the sociolinguistic process of focusing, which results in a reduction of the variants available from the pool of forms in the original mixture of dialects. Koineization enables the stabilization of the interdialect and gives rise to the stable, new-born dialect, or koine. In the new dialect, surviving variants from the original pool can potentially be reallocated as new sociolinguistic variants that are associated with social, stylistic, or areal meanings.

Trudgill (1986) also recognizes the importance of non-linguistic factors in the formation and development of new dialects. For instance, regarding which features from the original pool of variants will be eliminated through levelling, he claims that, "determining who accommodates to whom, and which forms are therefore lost, demographic factors involving proportions of different dialect speakers present will clearly be vital" (126).

However, as Tuten (2003: 29) points out, Trudgill "explicitly equates koineization only with the 'more purely linguistic forces' (Trudgill 1986: 126) of levelling and simplification."[60] Later studies such as Kerswill and Williams's (2000) investigation of the new variety of English in the Town of Milton Keynes have shown a critical role played by demographic and other factors during koineization. Such factors include the portion of adults vs. children in early waves of immigration, social network, literacy, etc.

---

[60] In fact, Tuten (2003) concludes that the processes of koineization also include "reallocation" as in Trudgill's terms. This, however, might not reflect Trudgill's real intention, as we already discussed above. Therefore, I exclude reallocation from the original quotation.

Therefore, in the current discussion of Dapeng formation, socio-historical factors have to be considered (§5.2-3).

5.2   Migration and Chinese Dialect Formation

As mentioned earlier, Siegel (1985) distinguishes two types of koines: regional koine and immigrant koine. The second category, "immigrant koine," is relevant in the case of Dapeng, as this dialect emerged in a new settlement established by migrants. Hakka and Cantonese are the source dialects of the Dapeng dialect, fulfilling the second criterion in Siegel's (1985: 365) definition of immigrant koine:

> "Two or more different linguistic varieties may be considered subsystems of the same linguistic system if they are genetically closely related and thus typologically similar enough to fulfill at least one of two criteria: (1) they are mutually intelligible or (2) they share a superposed, genetically related linguistic system, such as a national standard or literary language."

That is, although Hakka and Cantonese are overall not mutually intelligible, they do share both a national standard language and a literary language. Driven by immigration to the Dapeng area, as will be discussed in §5.3, Hakka and Cantonese have together given birth to the Dapeng dialect.

In fact, the issue of migration is of particular importance and relevance to the study of Chinese dialectology. The Chinese dialects did not develop and change in a vacuum. During the whole process of their development, socio-historical factors have played critical roles in shaping these dialects. Migration has been one of the major factors

124

that influence the formation and development of Chinese dialects, especially the Southern ones.

5.2.1 Models of Migration in the Chinese History

Migration has rarely ceased during the long history of China; after the first unification of China in the Qin Dynasty (approximately 200 B.C.), ancient Chinese people, wave after wave, moved around the country. Although the scale of migration has always varied, the general trends were from the North to the South and then from the East to the West (Li 2007: 20-21). Despite some cases of spontaneous migration, generally speaking, the Chinese demographic mobility in history can be described in one of the three following models (Ho 2015: 151-152):

1. The resettlement model (徙民模式 *xímín móshì*): motivated by the government for the purposes of opening up primitive regions, guarding the frontier, exiling criminals, etc.
2. The refugee model (流民模式 *liúmín móshì*): people fleeing their registered residence in times of war or natural disasters like famine and drought.
3. The invasion model (入侵模式 *rùqīn móshì*): driven by the invasions of the northern and western non-Han ethnic groups, which brought about ethnic and linguistic mixture.

According to Ho, while the resettlement model functioned mainly in very early eras and the invasion model contributed primarily to the formation and development of Northern Chinese dialects, the refugee model is the most relevant to the study of Southern dialects. He further distinguishes the two diverse situations that the refugee model involves regarding (Southern) dialect formation.

In cases where the destination of migration had been occupied by other Han Chinese people, the newcomers would have to settle in discarded or undeveloped lands surrounding the old communities, bringing about dialect contact mainly with the host group who lived nearby. In other cases when the original residents were non-Chinese people, the newcomers might live among the host group in the same area, which facilitated close contact between the transplanted Han Chinese and the indigenous non-Chinese language.

As centuries went by, the Han Chinese dialect usually ended up replacing its non-Han competitors. By then, however, elements of the non-Han languages had most likely permeated the Han Chinese dialect, leaving in it a substratum. Norman (1988) and Li (2007) both claim that this is the major process of the early foundation of the Southern Chinese dialects, out of which the modern Yue, Hakka, and Min dialects evolved. There are traces left on different aspects of modern Southern dialects from inter-ethnic contact.[61]

Based on these Non-Chinese substrata, Li (2007: 20-21) argues that subsequent waves of Han Chinese immigrations from the North entered into those Southern areas, each bringing in new dialects from their different origins and in different eras, adding new features to the local dialect. These features then formed newer strata in the local

---

[61] For instance, Norman (1988: 18, 213) proposes that the colloquial word for "shaman" or "spirit healer" shares similar forms along the southeastern coast, which gives clear indication of the Austroasiatic substratum under modern Southern dialects: *tøiŋ* in Fuzhou Min 福州閩語, *thuŋ kuŋ tsai* in Zhongshan Yue 中山粵語, and *thuŋ sin* in Hakka. This is also the case in phonology: the implosives [ɓ-] and [ɗ-] in a few Min, Yue, and Hakka dialects (e.g. Wenchang Min 文昌閩語 in Hainan) reflect an early Tai-Kadai substratum (Ho 2015: 156).

dialect.[62] In short, it is evident that only after multiple waves of immigration and through

centuries of contact did the modern Min, Cantonese, and Hakka dialects spoken in

Southern China come into being. Taking into consideration the factor of migration

complicates the investigation of Southern dialects. As shown above, however, its

inclusion enables us to see a fuller picture of the formation and development of Southern

dialects.

5.2.2   Cases of Dialect Formation Induced by Migration

In several cases of dialect contact in China induced by massive migration, the

framework of koineization has been used to explain the formation process of new dialects

as outcomes of such contact. For instance, Kuo (2005) investigates the socio-

demographic data and dialect use of the original Mandarin population who migrated to

Taiwan for political reasons after World War II in the mid-1940's. She argues that the

eradication of the retroflex initials /tʂ/, /tʂʰ/, and /ʂ/ was highly determined by the

demographic composition of that group. The retroflex initials were used by a relatively

small number of Mandarin migrants, while the majority of the migration group used

alveolar initials /ts/, /tsʰ/, and /s/ only. Therefore, the retroflex initials were more easily

---

[62] For instance, Cantonese shows substrata in all its lexicon and structure. Cantonese features mentioned in different parts of this dissertation in fact reveal different historical strata. A group of colloquial lexical items (such as "shaman") suggests the earliest stratum, which dates back to the early, indigenous non-Chinese language (this section); the phonological categories in general, as Norman (1988: 221) believes, were derived from Late Middle Chinese in the late Tang Dynasty (approximately the 9th and 10th centuries A.D., §2.2); the phenomenon of ditaxia (Matthew 1996), in comparison, should be ascribed to a very recent influence of modern Mandarin (§2.3).

levelled out in the competition of variants, and the alveolars survived in Taiwanese Mandarin.

In the Mainland Chinese context, Sun (2012) studies the Jianghan Oilfield Dialect (江汉油田话 *jiānghàn yóutián huà*), the *lingua franca* in the Jianghan Oilfield speech community. The rapid development of the petroleum industry in the 1960's and 1970's brought migrant workers to the rural area of Qianjiang 潜江, Hubei Province. Those workers were from many different Mandarin-speaking areas across Northern and Western China. Sun's study shows how today's Jianghan Oilfield Dialect has been formed both by levelling out the differences in the input Mandarin dialects and by further simplifying the structural complexity in the koine. Both the first and second generations of immigrants were involved in these processes, which have overtime stabilized the Jianghan Oilfield Dialect. Now, the koine has become the mother tongue of the third generation.

A similar koine-forming process is reported by Yang (2013). According to his study of the Shangrao Railway Dialect (上饒鐵路話 *shàngráo tiělù huà*) in Jiangxi Province, the residential community of the railway industry in Shangrao was established by hundreds of migrant workers from Zhejiang Province, mostly speakers of different Wu dialects. Starting in the 1940's, it also took two generations for the Shangrao Railway Dialect to emerge and develop through levelling and simplification of those Wu dialects. The koine had become stabilized by the end of the second generation. Different from the Jianghan Oilfield Dialect, however, the Shangrao Railway Dialect did not become the mother tongue of the third generation, where the speakers have generally shifted to Putonghua.

Given previous studies that have demonstrated the importance of studying socio-demographic context, a discussion of the formation process of the Dapeng dialect in this chapter will begin with a close examination of the immigration history of Dapeng (§5.3). As we will see, the socio-historical setting of the Dapeng dialect formation falls precisely into the category of Siegel's (1985) category of "immigrant koine." The hypothesis of koineization will then be further supported by linguistic evidence (§5.4).

While previous studies focus mainly on Northern and Central Chinese dialects, with Kuo (2005) and Sun (2012) on Mandarin and Yang (2013) on Wu, the current study investigates the Dapeng dialect as an outcome of Southern Chinese dialect contact between Cantonese and Hakka. Also, while previous studies discuss relatively new dialects formed no earlier than the 1940's, the current study involves much greater time depth, in the emergence of the Dapeng dialect several hundred years ago. From both geographical and historical perspectives, this study brings new contributions to our knowledge of contact-induced dialect formation in the Chinese context.

5.3   Socio-historic Evidence for Koineization

5.3.1   The Earliest Settlers (14th Century)

According to the local annals and other historical records (Baoan County Annals Committee 1997, Shenzhen Bowuguan 1997, Ji 2001, Zhang 2006, etc.), Dapeng began life as one of the military fortresses along the South China Sea. Built in 1394, its purpose of building this fortress was to serve as a stronghold against frequent attacks from pirates and foreign invaders.

According to Cao (1997: 317), it remains unclear whether the border guards in Guangdong Province included local soldiers, but judging from the situation in other provinces such as Zhejiang at that time in the early Ming Dynasty, soldiers serving the Guangdong costal fortresses were most likely locals. The early troops in Dapeng were also most likely soldiers recruited locally from Guangdong Province. Therefore, it is reasonable to surmise that those early settlers and troops in 14th century Dapeng, spoke different Chinese dialects of Guangdong, probably some early variation of Cantonese or Hakka dialects.[63]

Encouraged by the central government, immigrant soldiers brought their families, permanently settling in the Dapeng peninsula area. The first wave of immigrant troops included 3 officers and 133 soldiers (Baoan County Annals Committee 1997: 574). The number of people in the families was not recorded.

To accommodate the needs of the early troops, villages with new settlements gradually surrounded the military fortress. The soldiers rotated their duties between guarding the fortress and farming in the surrounding fields, a Ming policy that ensured that part of the burden of feeding the empire rested on the soldiers themselves (Wang 1965).[64] Due to frequent interaction both within the fortress and in the surrounding farming areas during the years, over time the soldiers and their family members naturally

---

[63] According to Hashimoto's (1973: 4) account of the five major waves of Hakka migration, there had been Hakka migrants in northern and eastern parts of Guangdong since the beginning of the 12th century (end of Song Dynasty).

[64] This phenomenon is recorded in the history records as "三分守城, 七分屯种 *Sānfēn shǒuchéng, qīfēn chúnzhòng,*" literally meaning "thirty percent (of the time or duty) guarding the fortress and seventy percent farming in the fields."

created a "common language" to facilitate communication within and outside the Dapeng fortress.

5.3.2 The Great Evacuation and Re-immigration (17th Century and Onward)

After the settlement in the late 14th century, the earliest troops and their descendants lived on the Dapeng peninsula for the next two centuries. This ended abruptly in 1661 when the central (Qing) government decided to evacuate a large area along South China Coast in order to cut off any possible material supplies to pirates, foreign invaders, as well as the armed navy force that was attacking intermittently from Taiwan.

Under such political pressure, all civilians in the villages surrounding Dapeng had no choice but to flee. They were all forced to move inland for tens of miles, while the entire Southeastern coast became a restricted military zone. Any civilians crossing the erected border walls or engaging in illegal trading would be severely punished, including that of beheading. This was called the Great Evacuation (遷界 *qiānjiè* or 遷海 *qiānhǎi*) of the Qing Dynasty (Baoan County Annals Committee 1997, Shenzhen Bowuguan 1997, Guangdong Province Annals Committee 1999, etc.).

For example, in the case of Xin'an County 新安縣, which included the current territory of Shenzhen, more than 90% of the land was abandoned and the vast majority of the local population was evacuated (Tan 2010: 224). In the same period, historical records show that by roughly 1668 about 800 soldiers were left in the Dapeng fortress to

keep watch of the South China Sea for pirate activities (Baoan County Annals Committee 1997: 574).

Approximately a decade after the evacuation, when civilians were again allowed to resettle in the Dapeng peninsula area in 1670, the majority of the original residents who had moved away did not return. Instead, hundreds of Hakka people from northern and eastern counties of Guangdong readily started moving in. They had been looking for a new place to inhabit for years. This demographic wave is also a branch of the fourth wave of the historical Hakka migrations (Hashimoto 1973: 4).

By 1688, the total number of Hakka households reached 230, with the total Hakka population at approximately 700 (Cao 1997, Tan 2010, Zeng 2011), and if the garrisoned troops are included in the count, then the total population of Dapeng doubles to approximately 1,500. In other words, by the end of the 17th century, the ratio of coastal guards who never left Dapeng during the Great Evacuation and the new Hakka immigrants was roughly 1:1.

Civilian immigration continued afterwards, but no subsequent large scale waves were recorded. In the meantime, the military population was slowly reduced in the 18th and 19th centuries. For instance, there were still 800 soldiers in 1726; in 1831, the number decreased to 505; in 1869, it was further reduced to 430. In 1899, the troops were withdrawn altogether from the Dapeng area, while civilians stayed (Yang and Huang 2001: 153-164).

In 2014, an unpublished government document showed that the population of permanent residents in Dapeng was approximately 3,000.[65] Most of these residents were native speakers of the Dapeng dialect. Based on the migration history of Dapeng, it is safe to infer that the current local population mainly consists primarily of descendants of two groups of residents co-existing there: soldiers and military officers of the Ming and Qing Dynasties, and the Hakka immigrants who arrived in the late 17th century.

5.3.3 The Formation of the Dapeng Dialect

The socio-historical background and demographic changes in Dapeng suggest that Trudgill's (1986) model of koineization may be able to account for the process of Dapeng's dialect formation. In close correspondence to the illustration in Figure 5 in §5.1.3, the formation steps of the Dapeng dialect are now hypothesized as follows.

First, some sort of common speech, a mixture of Cantonese and Hakka, should have emerged between 1400, the first period of recorded settlement and the Great Evacuation of the 1660's. Even after all civilians were forced to leave the peninsula during the Great Evacuation, the remaining soldiers would still have spoken that mixed dialect, which would have been more Cantonese like. Then, as new Hakka immigrants started to move in a decade later, the Dapeng population almost doubled within twenty years. This was the first encounter of the two groups and their dialects.

By 1688, approximately half of the local population consisted of Hakka speakers (700 out of 1,500). They brought new features to the pool of linguistic variants as they

---

[65] A more detailed demographic description of current-day Dapeng is given in §6.4.1.

settled into the peninsula community. Mutual accommodation would have taken place as the two groups of people (the remaining soldiers and the new immigrants) interacted with each other in daily life. After the first encounter, accommodation between both dialect groups began to take place.

It can also be inferred that in a presumably fairly short period of time, a new, independent identity tied specifically to the new Dapeng community would emerge.[66] Frequent communication within such a geographically closed area and the urgent need for cooperation for efficient coastal defense could also accelerate the development of a common identity. At this point, koineization began to take place, a topic to be discussed in §5.4.

5.4   Linguistic Evidence for Koineization

The previous section has examined the demographic evidence and hypothesized the Dapeng dialect formation as a case of koineization from a socio-historical perspective. This section examines linguistic evidence and discusses how well the linguistic processes of levelling and simplification, both leading to the reduction of variants in the original mixture of dialects, are able to account for the hypothesis. In this section evidence is drawn from various aspects of the Dapeng linguistic structure, using the dialect descriptions given in Chapters 3 and 4.

---

[66] This could happen in a generation or two, based on Kerswill and Williams' (2000) observation of the role of children in dialect levelling

5.4.1 Finals: Synchronic Perspective

One piece of evidence is found in the comparison among the final inventories of Dapeng, Hakka, and Cantonese, as shown in Figure 6. The figure shows that the final inventory of Dapeng is notably smaller than the two major Chinese dialects, both Hakka and Cantonese: the Dapeng dialect only has 41 finals, which is in sharp contrast to Hakka and Cantonese, with 76 and 68 finals, respectively.

Dapeng

a		ɔ	i	u
a				
ai	ɐi			ui
au	ɐu		iu	
am	ɐm		im	
an	ɐn		in	un
aŋ	ɐŋ	ɔŋ		uŋ
ap	ɐp		ip	
at	ɐt		it	ut
ak	ɐk	ɔk	ik	uk
ia	iɛ	iɔ		
iaŋ		iɔŋ		
iak		iɔk		
				m̩

Hakka

a	ɛ		ɔ	i	u
a					
ia	iɛ		iɔ		iu
ua	uɛ		uɔ		
ai			ɔi		
iai					ui
uai					iui
au	ɛu				
iau					
an	ɛn	ən	ɔn	in	un
ian	iɛn		iɔn		iun
uan	uɛn		uɔn		
am	ɛm	əm		im	
iam					
aŋ	ɛŋ		ɔŋ		uŋ
iaŋ			iɔŋ		iuŋ
uaŋ			uɔŋ		
ap	ɛp	əp		ip	
iap					
at	ɛt	ət	ɔt	it	ut
iat	iɛt		iɔt		iut
uat	uɛt		uɔt		
ak			ɔk		uk
iak			iɔk		iuk
uak			uɔk		
m̩					
ŋ̩					

Cantonese

a		ɛ	ɔ	œ	i	u	y
a							
ua			uɔ				
ai	ɐi	ei	ɔi	øy		ui	
uai	uɐi						
au	ɐu		ou			iu	
an	ɐn		ɔn	øn	in	un	yn
uan	uɐn						
am	ɐm				im		
aŋ	ɐŋ	ɛŋ	ɔŋ	œŋ	iŋ	uŋ	
uaŋ	uɐŋ		uɔŋ		uiŋ		
ap	ɐp				ip		
at	ɐt		ɔt	øt	it	ut	yt
uat	uɐt						
ak	ɐk	ɛk	ɔk	œk	ik	uk	
uak			uɔk		uik		
m̩							
ŋ̩							

41 finals   76 finals   68 finals

Figure 6. The Finals in Hakka, Dapeng, and Cantonese

Among many well studied Southern dialects, Hakka is special for the relatively large number of finals with the /-i-/ medial; Cantonese, in comparison, does not have the /-i-/ medial at all.[67] The Dapeng dialect has the /-i-/ medial, but its use is quite restricted.

---

[67] This is graphically suggested by the "longer" shape of the Hakka final inventory, as a large number of finals have their counterparts with the medial [-i-], which conventionally take up separate rows from the finals without [-i-]. For instance, towards the bottom of the Hakka inventory, [iak], [iɔk], and [iuk] are listed underneath [ak], [ɔk], and [uk].

136

On the other hand, one of the characteristics of Cantonese is its rich nuclei, which consists of eight main vowels in the phonological system, namely, /a/, /ɐ/, /ɛ/, /ɔ/, /œ/, /i/, /u/, and /y/. Hakka, in contrast, only has six: /a/, /ɛ/, /ə/, /ɔ/, /i/, and /u/. The Dapeng dialect has even fewer vowels: /a/, /ɐ/, /ɔ/, /i/, and /u/, with most of the shared finals between Cantonese and Hakka retained.[68] These contrasts suggest that the Dapeng final inventory is a somewhat reduced, simplified system, compared with Hakka and Cantonese. Neither salient features—the pervasive medial /-i-/ in Hakka or the rich nuclei in Cantonese—is favored in the Dapeng dialect.[69]

The observation based on the final inventories is also confirmed by Examples (27) and (28), with the former addressing the disfavor of the medial /-i-/ and with the latter illustrating the absence of /œ/. The salient differences of Hakka and Cantonese have been levelled out in dialect contact that gave rise to Dapeng, both replaced by more common variants that are shared by both input dialects.

---

[68] It should be noted again that in the Dapeng dialect there are six vowels phonetically but only five phonologically. The The phoneme /ɐ/ is realized phonetically as [ɛ] when it occurs after the medial [-i-], and as [ɐ] in all other instances.

[69] By "salient" I refer to features that make either Hakka or Cantonese distinctive from other southern Chinese dialects.

(27)

Dapeng	Hakka	Cantonese	Morpheme
mɐu^{54}	miau^{52}	mɐu^{22}	茂 "luxuriant"
kan^{42}	kian^{44}	kan^{53}	奸 "treacherous"
luk^{42}	liuk^{5}	lʊk^{2}	綠 "green"
hɐŋ^{42}	hiuŋ^{44}	hiŋ^{53}	兄 "older brother"
k^{h}ɐu^{22}	k^{h}iu^{44}	k^{h}ɐu^{23}	舅 "mother's brother"

(28)

Dapeng	Hakka	Cantonese	Morpheme
hiɔ^{42}	hiɔ^{44}	hœ^{53}	靴 "boot"
k^{h}i^{22}	k^{h}i^{44}	køy^{22}	巨 "huge"
tsɐn^{22}	tsun^{52}	tʃøn^{33}	俊 "handsome"
iɔŋ^{31}	iɔŋ^{11}	jœŋ^{21}	陽 "sun"
siɔk^{42}	siɔk^{1}	sœk^{33}	削 "pare"

Two important observations can be made here. First, when encountering salient features from either input dialect that are somewhat "marked" at least among Southern dialects in Guangdong, the Dapeng dialect constantly avoids adopting those features and tends to adopt the corresponding variants from the other dialect, which are less salient. For instance, Example (27) shows cases where the Dapeng dialect disfavors the medial /-i-/ and takes the Cantonese variants; Example (28), on the other hand, illustrates the avoidance of the salient phoneme /œ/ in Cantonese altogether,[70] in which case Dapeng closely follows the Hakka variants. In fact, the Dapeng dialect has no front rounded vowels at all.

Another observation is the ranking of avoidance of salient features in the input dialects. Competition exists between the variants of different contributory dialects, and it

---

[70] [œ] and [ø] are allophones of /œ/ in Cantonese.

is evident that the Dapeng dialect does not treat all salient features equally. In the case of 靴 "boot" and 削 "pare" from Example (28), neither input variant is favored, the Hakka form having the medial /-i-/ and its Cantonese counterpart having the phoneme /œ/. In situations like this, the choice of Dapeng is clear: the less salient variant is chosen.

The medial /-i-/, apart from thriving in Hakka, also exists in other Southern dialects, for instance, Southern Min spoken in Chaozhou 潮州, Guangdong (Beijing Daxue 2003: 35-37). The phoneme /œ/, on the other hand, is unique for Cantonese, given that it is the only dialect group in Guangdong Province that has this phoneme as a main vowel, according to Beijing Daxue (2003). Even within Cantonese, some of its dialects do not have this phoneme, especially those in the Siyi 四邑 subgroup, such as Xinhui 新會 and Doumen 斗門 (Zhan 2002: 242).

Therefore, the medial /-i-/ in Hakka is not as salient as the phoneme /œ/ is in Cantonese and thus wins the competition in the emerging koine, by being chosen by the Dapeng dialect. In this regard, one could infer that the choice of variant forms is not so much to adopt the most favorable ones, but rather to avoid the most unfavorable ones, in which the undesirable, most marked candidates of variants are levelled out of the competition.

5.4.2   Finals: Diachronic Perspective

From a diachronic perspective, a comparison among Dapeng, Hakka, and Cantonese based on the Middle Chinese final category also suggests the Dapeng sound system is the result of levelling and simplification of the input dialects. This is illustrated

in Examples (29) and (30), pertaining to the Middle Chinese 梗 Gěng and 曾 Zēng final

groups, respectively. First look at Example (29):

(29) Modern reflexes of the Middle Chinese 梗 Gěng final group
(a)

Dapeng	Hakka	Cantonese	Morpheme
maŋ^{31}	maŋ^{11}	maŋ^{21}	盲 "blind"
pak^{42}	pak^{1}	pak^{33}	百 "hundred"
saŋ^{35}	sɛn^{31}	ʃaŋ^{35}	省 "province"
tsʰak^{54}	tsʰɛt^{5}	tʃak^{2}	宅 "house"
tsʰaŋ^{54}	tsʰaŋ^{52}	tʃɛŋ^{22}	鄭 "Zheng (surname)"
tsʰak^{42}	tsʰak^{1}	tʃʰɛk^{33}	尺 "ruler"

(b)

lɐŋ^{31}	laŋ^{11}	liŋ^{21}	零 "zero"
tɐŋ^{42}	tɛn^{44}	tiŋ^{53}	丁 "labor"
sɐŋ^{42}	sən^{44}	ʃiŋ^{53}	星 "star"
sɐk^{42}	sət^{1}	ʃik^{5}	釋 "release"
kɐŋ^{42}	kin^{44}	kiŋ^{53}	京 "capital"
pɐk^{42}	pit^{1}	pik^{5}	璧 "jade"

Apparently, the Middle Chinese 梗 Gěng final group can be divided into two

groups in contemporary Dapeng based on main vowels: Group (a) has /a/ and Group (b)

has /ɐ/. The distribution of main vowels in Hakka seems somewhat random, especially in

Group (b), where four vowels are all possible: /a/, /ɛ/, /ə/, and /i/; /a/ and /ɛ/, in the

meantime, also appear in Group (a). On the other hand, pronunciations in Cantonese also

shows a clear pattern: in Group (a) /a/ and /ɛ/ are the two main vowels, while in Group

(b) only /i/ is used.

140

In fact, /a/, /ɛ/, and /i/ are the only three unrounded vowels in Cantonese preceding the ending /-ŋ/, while Dapeng only allows /a/ and /ɐ/ in the same position. Knowing this, the correspondence between Dapeng and Cantonese becomes clear: based on the openness of the main vowel, the Dapeng dialect matches the Cantonese close vowel /i/ with a relatively close vowel /ɐ/ in the Middle Chinese 梗 Gěng final group, and matches the other two (more) open vowels /a/ and /ɛ/ in Cantonese with an open vowel /a/. That is, between the two vowel distribution patterns in the source dialects, the Dapeng dialect has evidently followed the more regular one in Cantonese and levelled out the more irregular, Hakka pattern.

A very similar strategy is seen in the Middle Chinese 曾 Zēng final group. Example (30) lists some representative morphemes from this group.

(30) Modern reflexes of the Middle Chinese 曾 Zēng final group
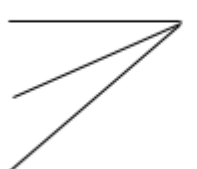
Dapeng	Hakka	Cantonese	Morpheme
jɐŋ^{42}	in^{44}	iŋ^{53}	鷹 "eagle"
lɐk^{54}	lit^5	lik^2	力 "strength"
tsɐŋ^{42}	tsən^{44}	tʃiŋ^{53}	蒸 "steam"
tsɐk^{42}	tsət^1	tʃik^5	織 "weave"
pɐŋ^{42}	pɛn^{44}	piŋ^{53}	冰 "ice"
sɐk^{42}	sɛt^1	ʃik^5	色 "color"
tɐŋ^{42}	tɛn^{44}	tɐŋ^{53}	燈 "lamp"
tɐk^{42}	tɛt^1	tɐk^5	德 "virtue"

As shown above, the pronunciation of the morphemes is again irregular in Cantonese and even more so in Hakka in terms of the choice of main vowel. The picture becomes more complicated if one considers the correspondence between these two

141

dialects, as there are at least four pairs of vowel correspondence: /i/-/i/, /ə/-/i/, /ɛ/-/i/, and /ɛ/-/ɐ/. The correspondences can be better illustrated in the chart in (31).

(31) Hakka-Cantonese vowel correspondences of the 曾 Zēng final group

Hakka   Cantonese

i ——————— i

ə

ɛ   ɐ

However, in the Dapeng dialect, one sees a consistent use of /ɐ/ as the main vowel in the Middle Chinese 曾 Zēng final group. Such consistency could have been achieved in two steps: first, the complicated distribution of vowels in both input dialects is eliminated, only leaving the shared vowel /i/; second, the Dapeng dialect matches the close vowel /i/ with a relatively close vowel /ɐ/ in its native phonology, which is allowed to proceed both the nasal ending /-ŋ/ and the stop ending /-k/.

Note that the Dapeng dialect only allows four main vowels to precede the nasal ending /-ŋ/: /a/, /ɐ/, /o/, and /u/. Only the first two are both [-rounded] and [-back], which match the feature of /i/. While /ɐ/ is not a perfect match of /i/ in terms of absolute closeness, compared with the open vowel /a/ in the Dapeng sound system, from a relative perspective /ɐ/ is nonetheless the closest vowel that could match /i/ from the source dialects. This match is more plausible if we consider the correspondence between the Dapeng /ɐ/ and the Cantonese /i/ from Example (29) in the Middle Chinese 梗 Gěng final group.

142

From the vowel correspondences in both Middle Chinese 梗 Gěng and 曾 Zēng

final groups, it can be seen that the Dapeng dialect does not simply adopt a variant from

either source dialect; instead, there is levelling-out of cross-dialectal differences based on

salience. The Dapeng dialect further adapts the choice of variant based on its native

phonology, aiming for more regularity of the sound system; that is, aiming for the

simplification of an originally more complex system resulting from dialect mixture.

These two types of innovation correspond to the two processes of koineization, levelling

and simplification.

5.4.3   Initial-Tone Integration

As stated in §3.4.1, the most striking feature of the Dapeng initial system is that

the Dapeng dialect shares the development pattern of Middle Chinese voiced obstruents

with Hakka. In both Hakka and Dapeng, those ancient voiced obstruents have been

devoiced and became aspirated regardless of the tonal conditions. In Cantonese, on the

other hand, they have become aspirated in syllables with the Middle Chinese *Píng* 平

("level") tone and in the colloquial reading layer in the *Shǎng* 上 tone (§2.2.3); in the

other tonal conditions, the ancient voiced obstruents are now unaspirated.

Examples in (32) provide a list of representative morphemes belonging to the

initial category of Middle Chinese voiced obstruents. In addition to the apparent, overall

correspondence between Dapeng and Hakka in aspiration, a close examination of this list

reveals more linguistic evidence which suggests the koineization process of Dapeng

formation.

143

(32) Modern reflexes of Middle Chinese voiced obstruents

Dapeng	Hakka	Cantonese	Morpheme
kʰɐu^{22}	kʰiu^{44}	kʰɐu^{23}	舅 "mother's brother"
pʰau^{22}	pʰau^{52}	pʰou^{23}	抱 "hug"
pʰui^{22}	pʰi^{52}	pʰui^{23}	倍 "multiple times"
tsʰi^{22}	tsʰu^{44}	tʃʰy^{23}	柱 "pillar"
pʰɐŋ^{22}	puŋ^{31}	pʰaŋ^{23}	棒 "stick (n.)"
tsʰi^{22}	si^{52}	tʃʰi^{23}	似 "similar"
kʰin^{54}	kʰian^{52}	kin^{22}	件 "measure word for clothes"
kʰɐi^{54}	kʰui^{31}	kuɐi^{22}	跪 "kneel down"
tʰau^{54}	tʰau^{52}	tou^{22}	道 "road"
tsʰui^{54}	si^{52}	tʃøy^{22}	序 "order (n.)"
pʰak^{54}	pʰak^{5}	pak^{2}	白 "white"
tsʰuk^{54}	siuk^{1}	tʃʊk^{2}	俗 "vulgar"
pʰu^{54}	pʰu^{52}	pou^{22}	步 "step"
tʰai^{54}	tʰai^{52}	tai^{22}	大 "big"
tsʰi^{54}	tsʰu^{52}	tʃy^{22}	住 "live (v.)"

The first observation is that while the Dapeng dialect overall parallels Hakka in the initial aspiration, the Dapeng initials do not preserve the occasional irregularities in the original, inputting Hakka initial system. The "devoicing plus aspiration" rule of initial change of ancient voiced obstruents is sporadically violated in Hakka, several being shown in (32). For instance, unlike the rest of the morphemes with aspiration in Hakka, 棒 "stick (n.)" remains unaspirated, and 似 "similar", 序 "order (n.)", and 俗 "vulgar" have even changed the manner of articulation from affricate to fricative.

144

The Dapeng dialect, on the other hand, has spontaneously "corrected" such irregularity in the source dialect. Despite the inconsistency in today's Hakka, all Middle Chinese voiced obstruent initials are consistently aspirated in the contemporary Dapeng dialect. Such modification could be a result of the simplification process where complexity and irregularities are reduced. That is, after the adoption of the general Hakka feature, i.e. aspiration of Middle Chinese voiced obstruents in the levelling process, the irregular distribution of voiceless unaspirated obstruents in the Dapeng initial system is fixed.

Another observation pertains to the major pattern of the Dapeng dialect neatly matching the Cantonese tonal categories. It has been shown from Table 3 in §3.2.4 that there are five tones in the Dapeng dialect, two of which are 22 (low level) and 54 (high falling). Tracing from a diachronic perspective, the low level 22 tone was mainly derived from the Middle Chinese *Yīnqù* 陰去 tonal category, supplemented by a part of the *Yángshǎng* 陽上 tonal category, and the high falling 54 tone in Modern Dapeng was mainly derived from the Middle Chinese *Yángqù* 陽去 and *Yángrù* 陽入 tonal categories, supplemented by another part of the *Yángshǎng* 陽上 category. Both of these parts from the *Yángshǎng* 陽上 category evolved into either the modern 22 (low level) or the 54 (high falling) tone associated with voiced obstruent initials in Middle Chinese, which is called *Quánzhuó* 全濁 in traditional terminology.

The rest of the *Yángshǎng* 陽上 category, mainly involved with nasal initials (*Cìzhuó* 次濁), has been merged with the *Yīnshǎng* 陰上 category, becoming the high

145

rising tone (35) in the Dapeng dialect. It was stated in §3.4.3 that the Dapeng dialect does

not separate the *Shǎng* tone 上聲 into the *Yīn* 陰 and the *Yáng* 陽 registers, which is the

case synchronically. From a diachronic perspective, the *Yīn* 陰 register of the Middle

Chinese *Shǎng* tone 上聲 (the *Yīnshǎng* 陰上 tonal category) is preserved in Modern

Dapeng, while the Middle Chinese *Yángshǎng* 陽上 tonal category has been split into

three modern tones and merged with the *Yīnshǎng* 陰上, *Yīnqù* 陰去, and *Yángqù* 陽去

plus *Yángrù* 陽入 tonal categories, respectively. The direction of tonal change from the

Middle Chinese sound system to the present-day Dapeng dialect is shown in Chart (33).

(33) Modern reflexes of some Middle Chinese tones in Dapeng

Middle Chinese categories	Modern Dapeng categories
Yīnshǎng 陰上	Tone 3 (35, high rising)
Yángshǎng 陽上	
Yīnqù 陰去	Tone 4 (22, low level)
Yángqù 陽去 + Yángrù 陽入	Tone 5 (54, high falling)
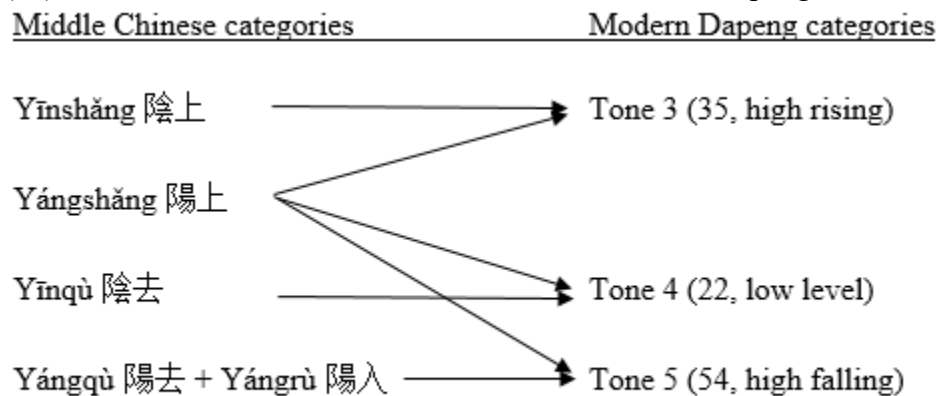
The evolution pattern of the Middle Chinese *Yángshǎng* 陽上 tonal category

associated with voiced obstruent initials has been an important cue for Chinese dialect

comparison, especially dialect grouping (Li and Xiang 2009). In the case of Dapeng,

Example (32) shows a neat match between Dapeng and Cantonese: all morphemes with

the tone value 23 in Cantonese and associated with aspirated initials have the low level

tone (22) in the Dapeng dialect, while all those with the tone value 22 or 2 in Cantonese[71] and associated with unaspirated initials have the high falling tone (54) in the Dapeng dialect. The same sort of clear match, by contrast, is not found between Dapeng and Hakka, nor between Cantonese and Hakka.

Compared with features in the final inventories (§5.4.1), it is more difficult to determine which pattern of distribution for the Middle Chinese *Yángshǎng* 陽上 tonal category is more "marked" in Cantonese vis-à-vis Hakka. Nonetheless, the tonal system of the Dapeng dialect has evidently adopted the more regular one between the two patterns of the input dialects, namely, Cantonese.

Although according to Trudgill's (1986: 126) definition, the "reduction of phonological and morphophonemic complexity and irregularities" mainly takes place in the simplification stage of koineization, in the case of the Middle Chinese *Yángshǎng* 陽上 tonal category, the distribution becomes obvious, as a newly formed dialect like the Dapeng dialect could also pursue structural regularity in the earlier, levelling stage. That is, when markedness is unclear, variants with more regularity from the input dialects would more likely be adopted. In other words, throughout koineization, regularity is the consistent goal for the new dialect in both processes of levelling and simplification.[72]

---

[71] In Cantonese the tones 22 and 2 have the same contour with different duration, as 2 associated with syllables ending with a stop ending, which are shorter than other syllables.

[72] As Hudson Kam and Newport (2005) demonstrate, processes of regularization in language contact situations (such as creolization) are particularly associated with children. With variable input from different contributory languages available, children most likely play the role of regularizing complex variability and imposing systematicity onto the new language as they learn it.

5.4.4   Multiple Layers of Pronunciation

As stated in §2.2.3, an important phenomenon among Southern Chinese dialects is the split between the literary and colloquial layers of phonological forms, which is mainly due to formal literacy education. Some morphemes therefore have differentiated pronunciations, again, with the one at the colloquial level more informal and everyday and the one at the literary layer more formal and elevated.

This somewhat complex pronunciation system exists in both Hakka and Cantonese but is absent in the Dapeng dialect. This difference across dialects is shown in (34), with the example morpheme again drawn from the Middle Chinese voiced obstruent initial category. (The first pronunciations of the doublets are from the colloquial layer, the latter from the literary layer.)

(34) Multiple layers of pronunciation in modern Chinese dialects

Dapeng	Hakka	Cantonese	Morpheme
$k^h\text{ɐn}^{22}$	$k^h\text{iun}^{44}$, $k^h\text{iun}^{52}$	$k^h\text{ɐn}^{23}$, kɐn^{22}	近 "close"
$t^h\text{in}^{22}$	$t^h\text{ɔn}^{44}$, $t^h\text{ɔn}^{52}/t^h\text{ɔn}^{31}$	$t^{h23}\text{yn}$, tyn^{22}	斷 "break (v.)"
$ts^h\text{ɔ}^{22}$	$ts^h\text{ɔ}^{44}$, $ts^h\text{ɔ}^{52}$	$t\int^{h23}\text{ɔ}$, $t\int\text{ɔ}^{22}$	坐 "sit"
$t^h\text{ɐi}^{54}$	$t^h\text{ai}^{44}$, $t^h\text{i}^{52}$	tɐi^{22}	弟 "younger brother"
$t^h\text{uŋ}^{54}$	$t^h\text{uŋ}^{44}$, $t^h\text{uŋ}^{52}$	tʊŋ^{22}	動 "move (v.)"

Example (34) shows that morphemes with two (occasionally three) layers of phonological forms in the two source dialects have only one corresponding layer in the Dapeng dialect. Also, the pronunciations in the Dapeng dialect consistently match those at the colloquial layer in Cantonese, as suggested by the alignment of tonal categories between Dapeng (22) and Cantonese (23).

This observation was based on the morpheme list, then confirmed by further investigation conducted during fieldwork. After noticing the seeming lack of differentiated layers of pronunciation, I compiled a list of 23 morphemes that normally have at least two phonological forms in Southern Chinese varieties. I then constructed proper context that when read would suggest different layers of pronunciation for these morphemes. One of my senior informants previously worked as an elementary Chinese language/literature teacher and was also a famous singer of local folklore. He was recommended to me for his rich knowledge of the local dialect. When he read the list, all of the morphemes had only one pronunciation—all from the colloquial layer. The result of this investigation supported my original impression from everyday observation and from the study of the comprehensive morpheme list.

Given the closeness in meaning and the relatively small number of morphemes that have multiple layers of phonological forms, it creates little problem, if any at all, if a Southern dialect eliminates one of those layers. Both the Hakka and Cantonese dialects preserve the multiple layers as a result of the long-term tradition of a formal education grooming students to participate in the central government workforce. The tradition stems back so long into Chinese history that even before the Dapeng dialect was formed, there had most likely been different layers of reading in both earlier forms of Hakka and Cantonese. We do not know why the Dapeng dialect has not created its own system of multiple-layer pronunciation; what is clear, nevertheless, is that the Dapeng dialect did not inherit such a system from either Hakka or Cantonese.

Of the two processes involved with the loss of multiple layers of pronunciation, simplification seems more relevant. Since those pronunciations normally did not differentiate meanings, the multiple layers of phonological forms became redundant. Redundancy inevitably brought in undesirable complexity to the dialect system. Given Dapeng's primary function as a military fortress where literacy was not as crucial as oral communication in general, redundancy in literary reading was even more likely to be reduced. When reduction was both favorable and plausible, simplification was naturally triggered, and irregularities decreased in the Dapeng dialect.[73]

5.4.5   Syntax

Both levelling and simplification are linguistic processes of koineization that occur in the formation of Dapeng. This can also be observed in the syntax of Dapeng. While the Dapeng syntax overall resembles both Hakka and Cantonese, especially so with the latter, the Dapeng dialect still has several syntactic features different from Cantonese and/or Hakka (§4.3). Two of these distinctive features of Dapeng also support the hypothesis of koineization.

First, in the disposal construction the Dapeng dialect uses the preposition 將 [tsiɔŋ^{22}] to introduce the affectee, a nominal that is influenced or disposed of by the action of the affecter. The basic structure is S + tsiɔŋ^{22} + O + V, meaning "the object is affected by the subject in the manner of the action V" (§4.3.3). The same morpheme is also used in both Hakka and Cantonese. In Hakka, there are several other disposal

---

[73] Even in Putonghua, the literary layer has overall been removed.

150

markers, such as 搭 [tap[1]], 捉 [tsɔk[1]], and 把 [pa[31]] (Li and Xiang 2009: 217), but 將 [tsiɔŋ[44]] (pronunciation in Dapeng) is the only morpheme it shares with Cantonese that could function as a disposal marker. In adopting this marker, the Dapeng dialect levelled out the differences between Hakka and Cantonese and only kept the common one. This is a result of levelling.

Second, as pointed out in §4.3.7, both Cantonese and Hakka make a distinction between progressive and continuous imperfective aspects. Cantonese distinguishes 緊 *gan[35]* (PROG) from 住 *zyu[22]* (CONT) (Zhan et al. 2002) and Hakka distinguishes 緊 *kin[24]* (PROG) from 等 *ten[24]* (CONT) (Song 2008). The marker in the Dapeng dialect does not have either differentiation of the source dialects; instead, the same post-verbal marker 緊 [kɐn[35]] plays the role of both the progressive and continuous imperfective markers. Another marker 到 [tau[22]] functions as an occasional alternative. Such reduction of complexity most likely emerged through the process of simplification, during which the Dapeng dialect lost the original distinctions in the perfective aspect that is present in the input dialects, Hakka and Cantonese.

5.5   Summary

This chapter relied on Trudgill's (1986) model of "koineization" and hypothesized that the formation process of the Dapeng dialect is koineization. Much of the discussion in the first half of the chapter was devoted to the clarification of terminology, arguing that koineization *per se* is not a unique linguistic process but is a

combination of two separate processes: levelling and simplification. The argument for both processes is supported in the second half of the chapter with linguistic evidence from the Dapeng dialect, making use of linguistic structures previously described in Chapters 3 and 4.

In the process of levelling, the Dapeng dialect strategically adopted features from either Hakka or Cantonese, the choice being dependent on the relative salience of the features in either source. In the process of simplification, the Dapeng dialect further adapted those features to achieve less complexity and higher regularity. The hypothesis of koineization was also supported by socio-historical evidence through a study of the migration history of the region.

Before this chapter ends, a note should be made about the comparison between the Dapeng dialect and its two contributory dialects, Hakka and Cantonese. Although in a "dialect report" it is conventional to compare synchronically the dialect in question with other contemporary dialects, for the purpose of hypothesizing the contact-induced formation process of a mixed dialect, it is justifiable to compare the earlier dialects, both the one in question and the contributing ones. It would be ideal if dialects spoken in the era of the initial contact could be directly compared.

Unfortunately, neither such data of Hakka nor the two input dialects is available to the current study. A few studies have reconstructed the sound system of earlier Cantonese, which date back to as early as the late 18th century (for instance, Peng 1992 and Zhao 2007). However, this date is still at least a century later than the initial contact of the early Hakka and early soldiers' speech in Dapeng around the 1680's (§5.3.3).

Given the lack of Hakka materials that belong to the same period, the reconstructed sound system of the late 18th century Cantonese cannot contribute much more to the cross-dialectal comparison than the contemporary Cantonese does. Considering the better quantity and quality of contemporary dialect data, this chapter compares contemporary dialects only. Albeit not ideal, this chapter has made the best educated hypothesis based on the data available.

Chapter 6   Assessment of Language Vitality

Building on the previous chapters, the current chapter turns now to the dialect speaking community and to assess the language vitality of the Dapeng dialect. It has been known that many local Chinese dialects in small communities are quickly losing their vitality in recent years. Given the overall similar linguistic ecology that the Dapeng dialect shares with those dialects, one would naturally wonder whether the Dapeng dialect is also losing its vitality and, if that is the case, at what pace is the vitality being lost? These questions will be addressed in this chapter.

This chapter will start with an overview of several representative studies of language vitality in Chinese dialectology (§6.1). It then goes on to review and compare some of the most influential frameworks that have been proposed on the assessment of language vitality (§6.2). Of particular interest and importance is the evaluation of whether and how much these models are applicable to the context of the Chinese language. I will argue that the UNESCO-LVE framework (Language Vitality and Endangerment scale, UNESCO Ad Hoc Expert Group on Endangered Languages 2003) is the most suitable one.

Then the UNESCO-LVE will be employed to assess the vitality of the Dapeng dialect (§6.3). Evidence comes from first-hand interviews, observations, and demographic data collected from my fieldwork. The results of assessment will show that

the overall vitality of the Dapeng dialect, although only spoken by a small population, is in fact fairly positive. This is in sharp contrast with many other small local dialects in China.

6.1   Language Vitality in Chinese Dialectology

It has been reported in recent decades that many local Chinese dialects in small communities spoken by limited people are gradually losing their vitality, with some examples being Zhanhua (站話, "the Post-station Speech") in Heilongjiang Province (You 1993), Chuanhua (船話, "the Fishermen's Speech") in Zhejiang Province (Cao 2001), Junhua (軍話, "the Army's Speech") in the Southeastern provinces (Huang 2007), and Shaoguan Tuhua (韶關土話, "the vernacular speech of Shaoguan") in Guangdong Province (Li and Zhuang 2009). All of these dialects are reported to be losing their vitality to varying degrees, some to the point of being severely endangered. For instance, Cao (2001: 10) predicts that Chuanhua may disappear within less than twenty years due to the impact of major dialects.

Huang (2007: 25) proposes four indicators of the loss of vitality in a Chinese dialect. These factors include: 1) a decreasing population of dialect speakers, 2) an aging population of dialect speakers, 3) increasingly restricted domains of use, and 4) the simplification of linguistic structure. Wu (2008) supplements this list with a few additional indicators, such as interrupted intergenerational transmission of the local dialect and negative language ideology. Based on the description provided in the reports

155

and surveys cited above, most of the local dialects mentioned above seem to share most of these symptoms.

Another obvious feature shared by these local Chinese dialects is that most of their native speakers are bilingual or multilingual. In addition to the local dialect, the speakers usually also speak some stronger, more prestigious Chinese dialects, be it Putonghua or the regional *lingua franca*. In the case of Shaoguan Tuhua, for instance, Tuhua speakers can most likely speak either—if not both—Cantonese or Hakka. It is thus believed that the growing impact from more prestigious, more powerful dialects is an external and yet crucial factor that accelerates local dialects' loss of vitality (Cao 2001, Sun 2001, Wu 2008, etc.).

The indicators proposed by Huang (2007) and Wu (2008) as well as bilingualism/multilingualism show an overall similar linguistic ecology shared by many minor dialects in China, including the Dapeng dialect. As the loss of vitality has been a prominent phenomenon among most of the small local dialects, one would naturally wonder whether the Dapeng dialect is also losing its vitality. If that is the case, at what pace is the loss in Dapeng, and at which stage is that loss? To answer these questions, an assessment of the Dapeng dialect vitality is conducted. The following section starts off with a brief review of the most influential frameworks that have been proposed in the literature of language vitality assessment.

6.2   Frameworks of Language Vitality Assessment

The endeavor of building frameworks to evaluate language vitality began in the early 1990's, when scholars started to propose scales to classify languages with different levels of vitality. Such frameworks include Schmidt (1990), Krauss (1992), Wurm (1998), etc.

These scales mainly focus on the generational transmission of language. Other aspects, such as domains of language use and language policies, are not sufficiently addressed. These scales are also relatively rudimentary, with only 4 to 5 levels of vitality.[74] On each level, the description of symptoms is usually short. In addition, there had been a "profusion of terms," as Florey (2005: 44) calls it, used to classify the levels of language vitality, including *safe, healthy, weakening, endangered, moribund, dying, extinct*, etc. It seems clear that some of the terms greatly overlap with each other, but the nuances in between are not specified.

Due to these shortcomings, these somewhat sketchy scales are not further discussed in this dissertation. The section below only focuses on three of the more comprehensive and mature frameworks, namely, Fishman's (1991) Graded Intergenerational Disruption Scales (GIDS), UNESCO's (2003) framework of Language

---

[74] Too few levels or categories could be problematic when assessing language vitality. For instance, based on the *Ethnologue* (1988, 11th edition), Krauss evaluates the vitality of about 6,000 languages and divides them into four categories: *safe, endangered, moribund*, and *extinct or dead*. He further predicts that only 10% of these languages could be called "safe," as he puts it: "I consider it a plausible calculation that—at the rate things are going—the coming century will see either the death or the doom of 90% of mankind's languages" (Krauss 1992: 7). Simons and Lewis (2013:8, 17) refer to Krauss' analyses as "necessarily sketchy and impressionistic" and "overly pessimistic", which was inevitably confined by the inadequate sources available to his time. In their opinion, from the perspective of the 21st century, language loss to such a degree as Krauss estimates will most likely not happen in regions other than Northern America and Australia, the areas where Krauss was most familiar with. (Also see Florey 2005: 43, Tsunoda 2005: 10, and Simons and Lewis 2013.)

Vitality and Endangerment (UNESCO-LVE), as well as Lewis and Simons' (2010)

Expanded GIDS.

6.2.1   Graded Intergenerational Disruption Scales (GIDS)

Fishman (1991) developed the Graded Intergenerational Disruption Scales

(GIDS) as an 8-level scale to describe the status of intergenerational transmission of a

given language. In his scheme, languages in a safer situation are denoted by a smaller

number. Languages classified at the six higher levels (Levels 1~6) are regarded overall as

safer, whereas the remaining two lower levels are on the more endangered end. A

language has less function and more restriction in its domains of use as the level

increases. At each level, as Table 9 shows, the GIDS explicitly states the features in the

speech community in terms of how languages are transmitted and used in particular

domains.

GIDS	(Adapted from Fishman 1991)	
LEVEL		DESCRIPTION
Safe	1	The language is used in education, work, mass media, government at the nationwide level
	2	The language is used for local and regional mass media and governmental services
↑	3	The language is used for local and regional work by both insiders and outsiders
	4	Literacy in the language is transmitted through education
↓	5	The language is used orally by all generations and is effectively used in written form throughout the community
Threatened	6	The language is used orally by all generations and is being learned by children as their first language
	7	The child-bearing generation knows the language well enough to use it with their elders but is not transmitting it to their children
	8	The only remaining speakers of the language are members of the grandparent generation

Table 9. The GIDS[75]

As the term GIDS indicates, the primary purpose of proposing this scheme is to examine the degree of language shift vs. language maintenance in a community, which is indexed by the situation of intergenerational transmission in the society (that is, disrupted vs. continuous transmissions). Besides showing where a language is at in this continuum of intergenerational transmission (from Level 1, fully used by most users to Level 8, used by few users), the GIDS also offers a means of revitalizing a language from a relatively threatened situation to a safer one, regardless of its current level. That is, if a language is assessed and placed at Level X, then for the purpose of revitalization, language activists can work on pushing the use of this language towards Level (X - 1) by increasing its

---

[75] This table is from Lewis and Simons (2010:25) with some adjustments made.

domain of use to meet the characteristics of the target level. This effort can continue until a desirable level of function and domain of use is reached. This process, as the title of Fishman's (1991) book suggests, is called Reversing Language Shift. Language revitalization, therefore, is the context and ultimate purpose of the GIDS model.

Lewis and Simons (2010:5-8) point out several shortcomings of the GIDS. These range from static terms and incomplete descriptions of all possible levels to the under-estimation of institutions with regard to the function of transmission and lack of detailed categories at the lowest level. Lewis (2008:35) also comments on this framework's failure to provide a detailed set of factors—apart from descriptions for each level—for researchers to evaluate language vitality.

Nevertheless, the GIDS is overall a detailed and coherent framework that classifies levels of vitality among many other early frameworks in the 1990's.[76] That explains why it remains the groundbreaking, foundational, most cited classification model for evaluating the various status of language vitality (Lewis and Simons 2010:4; Obiero 2010: 203-205). It has also inspired subsequent investigations on language vitality assessment and revitalization, many of which were conducted as a revision and an improvement of the GIDS (UNESCO 2003, Florey 2007, Lewis 2009, Lewis and Simons 2010, etc.).[77]

---

[76] The endeavor of building frameworks to evaluate language vitality began in the 1990's, when scholars such as Schmidt (1990), Fishman (1991), Krauss (1992), and Wurm (1998) proposed scales to classify languages with different levels of vitality. These early classification schemes are in general sketchy and relatively rudimentary, except Fishman's (1991) framework.
[77] See Fishman 2002, Brenzinger et al. 2003, and Tsunoda 2005: 10-11 for further introduction and discussion on this framework.

6.2.2 UNESCO Language Vitality and Endangerment Scale

The Language Vitality and Endangerment (UNESCO-LVE) scale was developed by the UNESCO Experts Meeting on Safeguarding Endangered Languages (UNESCO Ad Hoc Expert Group on Endangered Languages 2003).[78] This is a framework that consists of six major evaluative factors of language vitality, with two factors assessing language attitudes and policies, and one factor evaluating the urgency for documentation. Each factor serves to evaluate one aspect of a speech community separately using a 6-point scale. In all scales, the value "5" consistently denotes the most positive situation and "0" is assigned to represent the least positive. Then, crucially, these nine factors *together* contribute to the characterization of a language's overall vitality. The nine factors of UNESCO-LVE are listed below in Table 10:

---

[78] This work has also been cited elsewhere as (Brenzinger et al. 2003). In the rest of this chapter, it is sometimes referred to as the UNESCO official guide.

Factor 1	Intergenerational Language Transmission (scale)
Factor 2	Absolute Number of Speakers (real numbers)
Factor 3	Proportion of Speakers within the Total Population (scale)
Factor 4	Trends in Existing Language Domains (scale)
Factor 5	Response to New Domains and Media (scale)
Factor 6	Materials for Language Education and Literacy (scale)
Factor 7	Governmental and Institutional Language Attitudes and Policies, including Official Status and Use (scale)
Factor 8	Community Members' Attitudes toward Their Own Language (scale)
Factor 9	Amount and Quality of Documentation (scale)

Table 10. Nine Factors of UNESCO-LVE

Similar to many other frameworks, the UNESCO-LVE framework pays extra attention to the first factor, i.e. "Intergenerational transmission." It establishes "Safe" and "Extinct" at the two ends of the continuum of language vitality, and in between contains four more levels of vitality status: "Unsafe", "Definitely endangered", "Severely endangered", and "Critically endangered." As Lewis and Simons (2010: 8-9) have recognized, the UNESCO-LVE framework has different focuses compared with Fishman's (1991) GIDS regarding "Intergenerational transmission." The GIDS is more detailed on the "safe" end and sketchier on the "threatened" end, with six levels for the former and only two levels for the latter. In contrast, the UNESCO-LVE evidently puts more emphasis on the "threatened" end, as it identifies five levels at the "threatened" end but contains only one level for the status of "safe" end.

Besides "Intergenerational transmission", the UNESCO-LVE framework also calls attention to the other eight factors. Given the complexity and diversity of languages

and their speakers, the UNESCO-LVE framework emphasizes the importance of considering *all* factors collectively in order to evaluate the very specific situation of each speech community, as it indicates in the UNESCO (2003) official guide:

> "*No single factor alone can be used to assess a language's vitality or its need for documentation.*" (Page 7, italics in original)

> "The vitality of languages varies widely depending on the different situations of speech communities. The needs for documentation also differ under varying conditions. **Languages cannot be assessed simply by adding the numbers**; we therefore suggest such simple addition *not be done*. Instead, the language vitality factors given above may be examined according to the purpose of the assessment. (Page 17, bold in original)

> *The Factor descriptions … are offered as guidelines. Each user should adapt these guidelines to the local context and to the specific purpose sought.*" (Page 17, italic in original)

The emphasis to consider all nine factors together is the greatest strength of this model. It points out that, "Intergenerational transmission" being the central index of language vitality, is the result of and has close correlation with many other factors. The examination of intergenerational transmission indicates *where* a language is in the vitality continuum, while the other factors help explain both *why* it is there and *how* it functions at that level of vitality. The latter, if any revitalization is to take place, cannot be ignored.

In this regard, the UNESCO-LVE framework seems to have a contribution to language revitalization as the GIDS does, but it covers more aspects that contribute to the maintenance or loss of language vitality than its forerunner. For each factor, the UNESCO-LVE framework also makes available a detailed description of each of the six scales for investigators to compare their target language with. This comprehensive guide

makes the terms used in factor and scale names less vague, and with the guide at hand, it is more feasible for language evaluators to apply this framework to a wide range of languages. Given these merits, the UNESCO-LVE framework has been applied to the evaluation of language vitality worldwide since its publication in 2003.[79]

Nevertheless, this framework still has some flaws of oversimplification. As Obiero (2010: 207-209) points out, most of the nine factors are somewhat problematic and hence more or less need some revision. For instance, the exact meaning of "speakers" in Factor 3 is ambiguous, not specifying whether they are L1 or L2 speakers, nor if they are monolingual or multilingual. Factor 8 is another example: as language attitudes are, in practice, difficult to evaluate since "they are hardly ever uniformly held across an entire population." Despite these inadequacies, Obiero agrees that most of these factors raise fundamental questions about language vitality of some kind.

6.2.3 Expanded GIDS

As stated above, the two most frequently cited frameworks, Fishman's (1991) GIDS and UNESCO-LVE (2003) have different foci. While the former is much more detailed in distinguishing levels on the "safe" end, the latter focuses more on the "threatened" end. In order to develop a more complete framework that focuses on both

---

[79] For instance, Lewis (2006: 28) assesses the vitality statuses of 100 languages from around the world with the help of this framework, and concludes that it is "an admirable effort to bring together the 'state of the art' in terms of language endangerment and represents a reasonable and feasible approach to the issues with suggestions regarding appropriate responses" which "provides not only a clear framework for assessment but also delineates a very useful research agenda for investigators of the world's languages that is based on a sound theoretical orientation to language maintenance and shift." See Minasyan 2001 for more examples of the application of the framework to languages in Australia, North and South America, Africa, and other parts of the world.

ends, Lewis and Simons (2010) combine GIDS and NESCO-LVE and form an amplified

evaluative scale. This new framework aligns the GIDS and the UNESCO-LVE schemes

and is essentially an expansion of the former. As a result, it is called the "Expanded

Graded Intergenerational Disruption Scales", or EGIDS.

Based on the 16th edition of the *Ethnologue* (Lewis 2009), Lewis and Simons

categorize the global languages using a 13-level scale. Such a continuum presents a more

comprehensive range of language vitality across the world. In the EGIDS, the upper

levels (Levels 0 through 6a) deal with languages with relatively uninterrupted

generational transmission, which are further classified based on different situations of

language use and institutional support, while all the lower levels (from Level 6b to Level

10) cope with languages that are somehow not fully transmitted from one generation to

another, with generational transmission *per se* remaining the central criterion for the

classification of vitality levels. The full scale is shown in Table 11.

Level	Label	Description	UNESCO
0	International	The language is widely used between nations in trade, knowledge exchange, and international policy.	Safe
1	National	The language is used in education, work, mass media, and government at the nationwide level.	Safe
2	Provincial	The language is used in education, work, mass media, and government within official administrative subdivisions of a nation.	Safe
3	Wider Communication	The language is widely used in work and mass media without official status to transcend language differences across a region.	Safe
4	Educational	The language is in vigorous oral use and this is reinforced by sustainable transmission of literacy in the language in formal education.	Safe
5	Developing	The language is vigorous and is being used in written form in parts of the community though literacy is not yet sustainable.	Safe
6a	Vigorous	The language is used orally by all generations and the situation is sustainable.	Safe
6b	Threatened	The language is still used orally within all generations but there is a significant threat to sustainability because at least one of the conditions for sustainable oral use is lacking.	Vulnerable
7	Shifting	The child-bearing generation can use the language among themselves but they do not normally transmit it to their children.	Definitely Endangered
8a	Moribund	The only remaining active speakers of the language are members of the grandparent generation.	Severely Endangered
8b	Nearly Extinct	The only remaining speakers of the language are elderly and have little opportunity to use the language.	Critically Endangered
9	Dormant	There are no fully proficient speakers, but some symbolic use remains as a reminder of heritage identity for an ethnic community.	Extinct
10	Extinct	No one retains a sense of ethnic identity associated with the language, even for symbolic purposes.	Extinct

Table 11. Expanded Graded Intergenerational Disruption Scale[80]

---

[80] This table is cited from Lewis and Simons (2010: 28) and Simons and Lewis (2013: 22)

With the EGIDS making more levels available and all levels clearly named, scholars can now better identify a more accurate vitality status of a given language within the context of the wider range of an updated vitality continuum. Another merit of containing these multiple levels is the potential for more unambiguous coding of language vitality status. Applying this framework, Simons and Lewis (2013: 17) find that 63% of all human languages worldwide are at or above Level 6a, the critical level of "safely maintaining in everyday oral use in their communities", while 32% are at or Level 6b, the safety line and are thus at various levels of vitality loss. The remaining 5% have completely died out.

The combination of the GIDS and the UNESCO-LVE have enabled scholars to assess languages with various vitality profiles, both on the "safe" end and on the "threatened" end. Simons and Lewis (2013: 8-9) further state that, "the EGIDS can serve as a tool that is feasible to use on a global scale and that provides a better level of granularity and precision than other options that have been developed to date." Overall, the EGIDS is a comprehensive, widely used model. When applied to the Chinese languages,[81] however, this framework appears to encounter some problems. Those complications will be discussed in the following section.

---

[81] The term "Chinese language(s)" refers to languages or dialects (these two terms are not specifically distinguished here) pertaining to the Sinitic language family, such as Cantonese, Hakka, Min, etc. It does not refer broadly to "languages that are spoken in China." This term is used in the same way throughout the rest of this chapter.

6.3   Applicability of EGIDS and UNESCO-LVE to the Chinese Language

This section further compares assessment frameworks. In addition to the strengths and weaknesses pointed out in the previous section, I aim to evaluate whether and how much these frameworks are applicable to the vitality assessment of the Chinese language. Since the EGIDS is essentially an elaborated version of the original GIDS, I will only focus on the EGIDS and the UNESCO-LVE models, and will argue that the latter better fits the Chinese context.

The EGIDS, albeit comprehensive, encounters two major problems when applied to the Chinese language. The first problematic subject is Lewis and Simons' (2010) definitions of Levels 4 and 5 of the spectrum. Level 4, labeled as "Educational", contains languages that are "in vigorous use, with standardization and literature being sustained **through a widespread system of institutionally supported education**", while Level 5, with "Developing" as the label, denotes languages that are "in vigorous use, **with literature in a standardized form being used by some** though this is not yet widespread or sustainable" (bold added). If one is to faithfully follow the level placement criteria of the EGIDS, languages that do not meet the standards (in bold) should not be classified as higher than either a Level 4 or 5. In mainland China, therefore, no language other than Mandarin can be classified as at or higher than these two levels due to the general absence of both "institutionally supported education" and "literature in a standardized form being used."

But this is apparently not the case—even the *Ethnologue* editors, who labeled the vitality status for each language based precisely on the EGIDS model, seem to recognize

168

this. Under the entry of "Yue Cantonese,"[82] for instance, the status is labeled as Level 2 (Provincial), with additional supporting comments as "*De facto* provincial language in Guangdong Province." Take Hakka as another example, which is labeled at Level 5 (Developing).[83] According to the descriptions given above, on the other hand, these two Chinese languages should not be placed at a level higher than 6a due to the lack of institutional support and literature/orthographies, let alone a Level 5 or 2. Such inconsistency in level assignment suggests that the EGIDS framework may not be best applicable to the non-Mandarin Chinese languages given the complication of institutional support and literature/orthographies issues in the Chinese context.

We turn to consider the other option, the UNESCO-LVE model. Compared with the EGIDS developers, the UNESCO experts do not explicitly pinpoint *the* vitality status of each language. Unlike the EGIDS, the UNESCO-LVE framework does not aim to provide a conclusion of vitality status but rather to offer a comprehensive description of factors, each addressing a particular aspect of language use. This framework requires that the interpretation of vitality status has to be made based on *all* factors collectively and

---

[82] http://www.ethnologue.com/language/yue. This page contains information of Yue Cantonese spoken in mainland China only. For varieties spoken elsewhere, one more click on the "Also spoken in" drop list is needed, where further information concerning Hong Kong, Macau, Singapore, and many other Cantonese speaking areas will be displayed, each with its particular status of vitality. For instance, in the former two areas the level is also 2 (Provincial), while in Singapore it is 5 (Developing). In this dissertation, we only focus on the variety that is spoken in mainland China.
[83] http://www.ethnologue.com/language/hak.What adds even more to the complication is that, somewhere else under the same entry "Hakka", its status is indicated as Level 3 (Wider Communication) in the cloud of all living languages (http://www.ethnologue.com/cloud/hak). The same inconsistency happens with Southern Min, which is labeled as a Level 6a (Vigorous, http://www.ethnologue.com/language/nan) but shows a vitality of Level 3 in the cloud (http://www.ethnologue.com/cloud/nan). Although what is shown in the cloud seems to better match our impression of the real vitality profiles of these languages, I still stay with whatever is labeled. Such inconsistency, if beyond a reasonable extent, might nevertheless reduce one's confidence in the *Ethnologue* data.

that the local context and the specific research purpose being sought have to be considered while assessing language vitality.

Also, the nine factors in the UNESCO-LVE framework are guidelines, not rules. Even "generational transmission", the pivotal factor which alone may decide the vitality status in many other analytical models, is only one of the many factors and unable to define the status in the UNESCO-LVE model. Therefore, the UNESCO-LVE framework is by definition more conservative, less risky, and thus more explanatory in *complicated* situations, such as those in the context of the Chinese languages.

The UNESCO-LVE framework allows for cross-linguistic comparison and has been applied in various language contexts. A good example is shown at the end of the official guide of the framework (Page 23),[84] where three Venezuelan indigenous languages are compared side-by-side based on the UNESCO-LVE framework (Page 23, also shown in Table 12).

---

[84] For the sake of brevity, "Page xx" is used to indicate "Page xx from the official guide (UNESCO Ad Hoc Expert Group on Endangered Languages 2003)." The same below.

Factors	Languages		
	Mapoyo	Kari'ña	Sanima
1. Intergenerational Language Transmission	0	2	5
2. Absolute Number of Speakers	(7)	650	2500
3. Proportion of Speakers within the Total Population	1	2	5
4. Trends in Existing Language Domains	0	2	5
5. Response to New Domains and Media	0	1	---
6. Materials for Language Education and Literacy	1	3	0
7. Governmental & Institutional Language Attitudes and Policies including Official Status and Use	5	5	5
8. Community Members' Attitudes toward Their Own Language	2	3	5
9. Amount and Quality of Documentation	1	3	1

Table 12. Vitality of Three Venezuelan Indigenous Languages

This form of comparison suggests that the accumulation of fieldwork data, if evaluated by the same model, will make it possible to compare a number of languages in juxtaposition. In the context of the Chinese language, moreover, the UNESCO-LVE framework has another advantage. Some of the factors, especially those related to school education and orthography, may have a similar influence on non-Mandarin dialects, as those dialects rarely receive institutional support or have orthographic system.

For instance, a small local dialect may be graded as 0 ("no orthography available to the community") on Factor 6. While this is an undesirable score which points to low vitality, however, since most of the Chinese languages are graded identically here, the low score becomes a shared, redundant value, and thus should not be of concern for cross-dialectal comparison. Evaluators in this scenario can focus on other factors for

comparison among these languages/dialects. As studies of Chinese dialect vitality have not made wide use of the UNESCO-LVE framework, the current dissertation will be an early application of this framework.

6.4   Assessing Language Vitality of the Dapeng Dialect

This section applies the UNESCO-LVE framework to assess the vitality of the Dapeng dialect. The evaluation will be organized according to the nine factors in the framework.

Under each factor, the situation pertaining to the Dapeng dialect and the Dapeng community will be discussed, and a 6-degree scale based on 0 to 5 will be assigned, where the value "5" indicates the most positive situation and "0" denotes the least. For each factor, the brief description of all six degrees of the scale from the original UNESCO guide (2003) will be provided as well, and the category to which Dapeng falls will be in bold. At the end of this section, all nine grades will be put together and summarized in the form of an evaluation report.

In the following sections, §6.3.1 addresses the six major factors of language use, §6.3.2 pertains to the two factors that assess language attitudes and policies, and §6.3.3 discusses the factor that evaluates the urgency for documentation. Finally, §6.3.4 combines all nine factors and evaluates the overall vitality of the Dapeng dialect.

6.4.1   Major Factors of Language Use

Factor 1: Intergenerational Language Transmission

"Intergenerational Language Transmission" is the most important factor for language vitality, as it indicates how well a language is being transmitted from one generation to another. The scale of this factor is shown in Table 13.

Degree of Endangerment	Grade	Speaker Population
Safe	5	The language is used by all ages, from children up.
Unsafe	**4**	**The language is used by some children in all domains; it is used by all children in limited domains.**
Definitively endangered	3	The language is used mostly by the parental generation and up.
Severely endangered	2	The language is used mostly by the grandparental generation and up.
Critically endangered	1	The language is used mostly by very few speakers, of great-grandparental generation.
Extinct	0	There exists no speaker.

Table 13. Factor 1: Intergenerational Language Transmission

In the case of the Dapeng dialect, the choice is either Level 5 or Level 4, as the dialect is being used by all ages in the community, including children. One day I was waiting for my informants in their living room, and their twins (about 4 or 5 years old) were playing next to me. I heard them speaking the Dapeng dialect to each other during the entire time. I also asked my other informants what dialect their children (or grandchildren) spoke, and the answer was typically "the Dapeng dialect." Only in cases

where one of the parents was not local person did the children speak another dialect, or sometimes Putonghua. The Dapeng dialect was consistently one of the primary choices.

According to my observation and the interview with native speakers, the Dapeng dialect is evidently used by all ages, from children up, which is the description of Level 5. However, Level 5 cannot be assigned to the Dapeng dialect. This is because the children's use of the local dialect does not cover the full range of domains, and in particular not the domain of school education. With Putonghua promoted as the official language for over half a century, it is now used in the school setting in Dapeng, and the local dialect is no longer used.

The pervasiveness of Putonghua is also due to immigration. Having several huge factories (each having thousands of laborers, most of which are non-local) surrounding the Dapeng community, the Dapeng K-12 system has now enrolled a large number of migrant workers' children. The local children, no matter how much Dapeng dialect they speak at home or in the local community, need to speak Putonghua with their peers, in and after class. Therefore, the local dialect is not used by all children in all domains. The ultimate decision is the assignment of a Level 4 (Unsafe).

In fact, in the explanation attached to the degree table, there is an additional category named: "Stable yet threatened." This category is graded as (5-), apparently inserted between the "Safe" and "Unsafe" categories in Table 13. The (5-) Level is described as follows: "The language is spoken in most contexts by all generations with unbroken intergenerational transmission, yet multilingualism in the native language and one or more dominant language(s) has usurped certain important communication

contexts. Note that multilingualism alone is not necessarily a threat to languages." This appears to be a closer depiction of the situation in Dapeng. But in order to closely match other factors in the 6-point scale, I ultimately stayed with assigning the Dapeng dialect a Level 4.

Factor 2: Absolute Number of Speakers

"Absolute Number of Speakers" is the only factor among the nine that does not use a scale. Showing the real number of speakers of a community, this factor does not have specific descriptions provided in correspondence to degrees of endangerment. According to an unpublished census document,[85] in 2013 there were 630 permanent households, in total 1,828 residents in the local community. In addition, there was another population totaling 3,775 people, registered as having either Hong Kong or Macau, or sometimes some other citizenship (and thus not counted as permanent residents, who are by definition Chinese citizens). This second group has more mobility, and many are seasonal residents.

Therefore, it is always difficult to calculate an exact number of the total population of Dapeng dialect speakers, which consists of both these groups. The local government employee that I was in touch with gave me an unofficial estimate though: the population of the Dapeng dialect speakers living in Dapeng is normally around 3,000.

According to the UNESCO official guide (Page 8), a small population is more vulnerable and more subject to change than larger ones. Therefore, despite the lack of

---

[85] I obtained this document from the local government during fieldwork.

scales, one can infer that the limited absolute number of speakers of the Dapeng dialect, by itself, can hardly guarantee a stable, continuous transmission of the local dialect in the long term. On the other hand, other factors also need to be considered.

Factor 3: Proportion of Speakers within the Total Population

This factor considers the proportion of speakers who speak the target language in relation to the total population. The scale of this factor is shown in Table 14.

Degree of Endangerment	Grade	Proportion of Speakers
Safe	5	All speak the language.
Unsafe	**4**	**Nearly all speak the language.**
Definitively endangered	3	A majority speak the language.
Severely endangered	2	A minority speak the language.
Critically endangered	1	Very few speak the language.
Extinct	0	None speak the language.

Table 14. Factor 3: Proportion of Speakers within the Total Population

The definitions of Levels 4 and 3—"nearly all" vs. "a majority"—sometimes make it difficult to distinguish the two levels from each other. In the case of Dapeng, it still seems clear that the local language should fall again into Level 4. As introduced in §2.4, I interacted with dozens of senior residents and more than ten younger residents under the age of 35 during fieldwork. Among these groups of people who identify themselves as Dapengers, I only met one person claiming to not speak the local dialect well.

176

That person is a 23-year-old guard, whom I came into contact with through my transcriber, Miss Y. When I asked if he would like to be one of my informants, he hesitated and said he "does not speak the local dialect well enough." He said he can speak the Dapeng dialect but does not know the vernacular words well, and ultimately declined my request for audio recording. Maybe he is just shy, or maybe he is indeed not fully capable in the local dialect; but even if we consider this young man as a counter example, there are still many others who do speak the Dapeng dialect (and probably speak it well). I interacted with dozens of local people whose ages ranged from 22 to 84 but had only come across one such case during my entire stay. Therefore, it seems reasonable to place the Dapeng dialect on Level 4 (Unsafe).

Factor 4: Trends in Existing Language Domains

This factor reflects how the target language is used in discourse domains and what functions it has. The scale of this factor is shown in Table 15.

Degree of Endangerment	Grade	Domains and Functions
Universal use	5	The language is used in all domains and for all functions
Multilingual parity	**4**	**Two or more languages may be used in most social domains and for most functions.**
Dwindling domains	3	The language is in home domains and for many functions, but the dominant language begins to penetrate even home domains.
Limited or formal domains	2	The language is used in limited social domains and for several functions
Highly limited domains	1	The language is used only in a very restricted domains and for a very few functions
Extinct	0	The language is not used in any domain and for any function.

Table 15. Factor 4: Trends in Existing Language Domains

The issue of language domains has been partly touched upon in previous sections, and the Dapeng dialect seems to match Level 4 (Multilingual parity) criteria for this factor. The description given for Level 4, however, is not very clear in Table 15. The official guide (Page 9) has a supplementary explanation for Level 4 that is more informative:

> "One or more dominant languages, rather than the language of the ethnolinguistic group, is/are the primary language(s) in most official domains: government, public offices, and educational institutions. The language in question, however, may well continue to be integral to a number of public domains, especially in **traditional religious institutions, local stores, and those places where members of the community socialize**. The coexistence of the dominant and non-dominant languages results in speakers' using each language for a different function (diglossia), whereby the non-dominant language is used in informal and home contexts and the dominant language is used in official and public contexts. Speakers may consider the dominant language to be the language of social and economic

opportunity. However, older members of the community may continue to use only their own minority language. Note that multilingualism, common throughout the world, does not necessarily lead to language loss." (bold added)

This description overall reflects my observations of the Dapeng community, with only a few exceptions—school and government, communicating with Mandarin speakers, the official language, Putonghua, is rarely heard among Dapeng speakers in conversation with each other. The two more prestigious regional dialects, Cantonese and Hakka, are also seldom used by the local people. Dapeng speakers will switch to one of these dialects only when their interlocutors are insufficiently capable in the local speech. This is especially so for older speakers: As long as the other party, even outsiders, are able to understand the Dapeng dialect to some extent, they will use the local speech. All in all, the Dapeng dialect maintains its status well as the socially dominant usage across various domains of use in the local community.

Factor 5: Response to New Domains and Media

Factors 5 and 6 indicate two areas where the Dapeng dialect—as well as many other small local dialects in China—is assigned an extremely low level of vitality. Factor 5 deals with how well a language is able to expand its scope of use to newly emerging domains, such as new forms of education, new work environments, and new media (broadcast, the Internet, etc.). The UNESCO official guide (Page 11) warns that, "If the communities do not meet the challenges of modernity with their language, it becomes increasingly irrelevant and stigmatized." The scale of Factor 5 is shown in Table 16.

Degree of Endangerment	Grade	New Domains and Media Accepted by the Endangered Language
Dynamic	5	The language is used in all new domains.
Robust/active	4	The language is used in most new domains.
Receptive	3	The language is used in many domains.
Coping	2	The language is used in some new domains.
Minimal	1	The language is used in only a few new domains.
Inactive	**0**	**The language is not used in any new domains.**

Table 16. Factor 5: Response to New Domains and Media

In a sharp contrast to the stable use in traditional domains (such as family, market, and farming, as indicated in Factor 4), the Dapeng dialect hardly shows any signs of vitality in the new domains and media. Similar to the situation discussed in Factor 4, all new domains examined in this study are entirely occupied by other major dialects, usually Putonghua, and sometimes Cantonese. Hence for Factor 5, the Dapeng dialect is placed at Level 0, that is, "inactive."

Factor 6: Materials for Language Education and Literacy

Factor 6, with the left-most column "Degree of Endangerment" missing, seems to be a relatively marginal factor, as it does not directly index degree of vitality. The scale of this factor is shown in Table 17.

Grade	Accessibility of Written Materials
5	There is an established orthography, literacy tradition with grammars, dictionaries, texts, literature, and everyday media. Writing in the language is used in administration and education.
4	Written materials exist, and at school, children are developing literacy in the language. Writing in the language is not used in administration.
3	Written materials exist and children may be exposed to the written form at school. Literacy is not promoted through print media.
2	Written materials exist, but they may only be useful for some members of the community; and for others, they may have a symbolic significance. Literacy education in the language is not a part of the school curriculum.
1	A practical orthography is known to the community and some material is being written.
0	**No orthography available to the community.**

Table 17. Factor 6: Materials for Language Education and Literacy

Local dialects' lack of orthography is common in the Chinese context, as non-Mandarin written material is generally lacking. There are perhaps two exceptions: first, major dialects, especially those spoken in wealthier regions where the vernacular literature is better developed (for instance, Cantonese in Guangzhou and Wu in Shanghai),[86] and second, dialects in which missionaries have translated the Bible and other religious scriptures for (for instance, the Cantonese, Hakka, and Min dialects along the coast). Dapeng falls into neither category, and is hence again assigned to Level 0, "No orthography available to the community."

---

[86] The Wu materials are also very limited.

6.4.2   Language Policy, Attitude, and Urgency for Documentation

Factor 7: Governmental and Institutional Language Attitudes and Policies, including

Official Status and Use

Factors 7 is related to language attitudes and policies, which are also relevant to
the evaluation of language vitality. It indexes the degree of governmental support
promoting dialect languages. The scale of this factor is shown in Table 18. In this scale,
both explicit policies and implicit attitudes from the government toward the examined
language are considered.

Degree of Support	Grade	Official Attitudes toward Language
Equal support	5	All languages are protected.
Differentiated support	4	Minority languages are protected primarily as the language of the private domains. The use of the language is prestigious.
Passive assimilation	**3**	**No explicit policy exists for minority languages; the dominant language prevails in the public domain.**
Active assimilation	2	Government encourages assimilation to the dominant language. There is no protection for minority languages.
Forced assimilation	1	The dominant language is the sole official language, while non-dominant languages are neither recognized nor protected.
Prohibition	0	Minority languages are prohibited.

Table 18. Factor 7: Governmental and Institutional Language Attitudes and
Policies, including Official Status and Use

First of all, in mainland China it is very rare for the Southern, non-Mandarin
dialects to receive governmental or institutional support of any kind. As a result, none of
them should be placed higher than Level 3. There might be argument, then, on whether a

Southern dialect is undergoing passive or active assimilation. Some might raise the example of the recent anti-Putonghua campaigns in Guangzhou; a similar type of tension, however, is not found in the Dapeng community.

In Dapeng, the local government is not aggressively promoting Putonghua, the official language. Compared with Cantonese, the Dapeng dialect is spoken by a significantly smaller population and has very limited influence outside the peninsula area. Therefore, it should not be seen as a threat of the official language. The local community, on the other end, is not taking any confrontational action to promote the Dapeng dialect beyond its old domains of use, either. According to my observation, the equilibrium is being well preserved. Hence, the Dapeng dialect is placed at Level 3, facing "passive assimilation" from the government.

Factor 8: Community Members' Attitudes toward Their Own Language

Factors 8 is also related to language attitudes from the community members' perspective. The scale of this factor is shown in Table 19.

Grade	Community Members' Attitudes toward Language
5	All members value their language and wish to see it promoted.
4	**Most members support language maintenance.**
3	Many members support language maintenance; others are indifferent or may even support language loss.
2	Some members support language maintenance; others are indifferent or may even support language loss.
1	Only a few members support language maintenance; others are indifferent or may even support language loss.
0	No one cares if the language is lost; all prefer to use a dominant language.

Table 19. Factor 8: Community Members' Attitudes toward Their Own Language

One of the most striking characteristics about the Dapeng speaking community that I found during my fieldwork is their loyalty to their mother tongue. Based on my previous knowledge about Shaoguan Tuhua, another small local dialect spoken in Guangdong Province, I arrived at Dapeng with an assumption of a similar generational difference in language attitude.[87]

However, the situation in Dapeng is significantly different. All people in the local community—literally everyone that I met—expressed their strong support toward the maintenance of the local dialect. Most of my interviewees, old and young alike, said with no hesitation that they have had, or will have their children speak the Dapeng dialect. Some even made the following claim: "One has to know how to speak the local dialect in order to qualify as a Dapenger." These responses are consistent to their preference of using the Dapeng dialect in most domains of language use, and their response and

---

[87] In the Shaoguan Tuhua community, only the older generation values their mother tongue (Li and Zhuang 2009, Chen 2012).

linguistic behavior both demonstrate a fairly positive attitude towards their own local dialect.

The only reason that the Dapeng dialect is placed at Level 4 instead of Level 5 is lack of evidence that the Dapeng speakers "wish to see the local dialect promoted." The Dapeng community is overall conservative, laid-back, and unambitious in (re)claiming domains of local dialect use beyond what they already have. In order to avoid overestimating the local language attitude, the Dapeng dialect speaking community is hence assigned a Level 4 for Factor 8.

Factor 9: Amount and Quality of Documentation

Factor 9 evaluates the urgency of documentation of a language. The scale of this factor is shown in Table 20.

Nature of Documentation	Grade	Language Documentation
Superlative	5	There are comprehensive grammars and dictionaries, extensive texts; constant flow of language materials. Abundant annotated high-quality audio and video recordings exist.
Good	4	There are one good grammar and a number of adequate grammars, dictionaries, texts, literature, and occasionally updated everyday media; adequate annotated high-quality audio and video recordings.
Fair	3	There may be an adequate grammar or sufficient amount of grammars, dictionaries, and texts, but no everyday media; audio and video recordings may exist in varying quality or degree of annotation.
Fragmentary	**2**	**There are some grammatical sketches, word-lists, and texts useful for limited linguistic research but with inadequate coverage. Audio and video recordings may exist in varying quality, with or without any annotation.**
Inadequate	1	Only a few grammatical sketches, short word-lists, and fragmentary texts. Audio and video recordings do not exist, are of unusable quality, or are completely un-annotated.
Undocumented	0	No material exists.

Table 20. Factor 9: Amount and Quality of Documentation

During my fieldwork, the only written materials that I found related to the Dapeng dialect were some lyrics of local ballads, which recorded some simple Dapeng grammar and a few colloquial words. Regarding audio- and video-recordings, there exist some clips of television reporters interviewing local residents speaking the Dapeng dialect. These are the only few documentations that were accessible to me. Based on this sparse materials, the Dapeng dialect is assigned to Level 2, "Fragmentary."

6.4.3   A Collective Evaluation of All Factors

Following the UNESCO-LVE model, I will now combine all of the nine factors together. The scales of all these factor shown in combination in Table 21.

Factors	The Dapeng dialect[88]
1. Intergenerational Language Transmission	4
2. Absolute Number of Speakers	3,000
3. Proportion of Speakers within the Total Population	4
4. Trends in Existing Language Domains	4
5. Response to New Domains and Media	0
6. Materials for Language Education and Literacy	0
7. Governmental & Institutional Language Attitudes and Policies including Official Status and Use	3
8. Community Members' Attitudes toward Their Own Language	4
9. Amount and Quality of Documentation	2

Table 21. A Combination of All Factors

As Table 21 indicates, in half of the first six factors, those which more directly link to language vitality, the Dapeng dialect is placed at the second highest level, namely Level 4. The two factors that show a tendency otherwise (viz. Factors 5 and 6) are those related to the common issues of inadequate institutional support and literature/orthographies shared by non-Mandarin Chinese languages, which was already addressed in Section 1.3. The levels that Dapeng assigned for Factors 5 and 6 are predictable as a Southern Chinese dialect. Based on Factors 1, 3, and 4, the overall

---

[88] More precisely, "the Dapeng dialect spoken in the Dapeng peninsula community."

vitality of the Dapeng dialect is in fact fairly positive. Considering the fact that the Dapeng dialect is such a small dialect spoken by about 3,000, which includes less than 2,000 permanent residents (according to Factor 2), the positive vitality of the Dapeng dialect is even more remarkable.

The evaluation based on the first six factors is further confirmed by the following two factors concerning language policies and attitudes. Although the Dapeng dialect, not unlike many other Chinese dialects, receives minimal or no support from the government, it has a strong base of its own speakers, who have very loyal, supportive attitudes towards the local dialect.

However, lack of documentation may become a problem, as Factor 9 shows. Therefore, an endeavor to further document the Dapeng dialect via various types of recording, be it textual, audio, or video should be made.[89]

6.5   Summary

This chapter discusses the issue of language vitality assessment of the Dapeng dialect. It first attempts to apply analytical frameworks to an unstudied Chinese dialect. This chapter carefully examines their applicability to the assessment of the Chinese language and proposes that the UNESCO Language Vitality and Endangerment (UNESCO-LVE) scale is most useful in the Chinese context. Then this framework is then applied to the evaluation of the Dapeng dialect vitality.

---

[89] The documentation of Dapeng, in fact, is exactly one of the main reasons why the fieldwork was conducted and why this dissertation is written.

Although some of the factors in the UNESCO-LVE scale suggest a fairly low vitality, an overview of all factors shows that the Dapeng dialect is still in a overall vigorous condition. Through the case of Dapeng, the results of evaluation suggest that not all small local dialects in China are necessarily in danger. It is true that a small population, lack of written material, pervasive bilingualism, among many other sociolinguistic factors, are all related to the loss of language vitality; however, this chapter has illustrated that none of these factors shall define the vitality status alone. Sociolinguistic factors are at work together; it is only by considering all factors collectively that one can reach a valid conclusion of the vitality profile of any particular language.

This chapter hence calls to attention the future studies of dialect vitality for Chinese related languages and dialects to a more comprehensive, systematic examination of sociolinguistic factors. This attempt, to the best of my knowledge, has not hitherto been sufficiently carried out in the field of Chinese dialectology.

Chapter 7   Conclusion

This chapter summarizes the major findings of this study in response to the research questions raised at the beginning of this dissertation (§7.1). It also discusses some limitations of this dissertation (§7.2) and offers some directions and suggestions for future studies (§7.3).

7.1   Summary of Findings

This dissertation has taken an initial step in describing and analyzing the hybrid nature of the Dapeng dialect as well as its use in the local community. It has three major findings, each corresponding to one of the research topic and questions from §1.1, as restated below:

1. Description of the Dapeng dialect

What exactly is the Dapeng dialect like? What are some of the linguistic features that distinguish it from other Southern Chinese dialects? In precisely what way does this local dialect show its hybrid nature of Hakka and Cantonese? From the perspective of its sound system, lexicon, and syntax, to what degree is it like Hakka? How does it resemble Cantonese?

2. Formation process of the Dapeng dialect

Based on a detailed description of the mixed nature of the Dapeng dialect, what are some possible formation processes of the Dapeng dialect one could infer? What theoretical framework(s) of dialect formation can be employed to account for the genesis of Dapeng? How well can such proposal(s) be supported by both linguistic evidence of the dialect and by socio-historical facts?

3. Language vitality of the Dapeng Dialect

Being a local dialect spoken by a small community in Southern China, how much vitality does the Dapeng dialect have in today's peninsula community? Is it endangered, as in the case of many other Chinese local dialects spoken in small communities under the influence of major, more "powerful" Chinese dialects? How do linguistic and social factors (such as bilingualism, language policy, and attitude) affect the maintenance and development of the Dapeng dialect?

7.1.1 Summary of Some Features of the Dapeng Dialect

First, this dissertation has provided a preliminary analysis of the Dapeng dialect. A detailed examination of the Dapeng sound system shows some degree of resemblance between the Dapeng dialect and the source dialects—both Cantonese and Hakka. The resemblance to either source dialect is intertwined at all aspects of the Dapeng sound system: initials, finals, and tones. For instance, in terms of the development of Middle Chinese voiced obstruents, the Dapeng dialect shares the pattern closely with Hakka, as shown in the initial correspondences in (35), originally (1) in §3.4.1.

(35)

Dapeng	Hakka	Cantonese	MC initial	Morpheme
p^hu^{54}	p^hu^{52}	pou^{22}	*b-	步 "step"
p^hu^{35}	p^hu^{31}	pou^{22}	*b-	捕 "to catch"
t^hau^{54}	t^hau^{52}	tou^{22}	*d-	道 "road"
t^hau^{22}	t^hau^{44}	tou^{22}	*d-	稻 "paddy"
pu^{22}	pu^{52}	pou^{33}	*p-	布 "cloth"
tau^{22}	tau^{52}	tou^{33}	*t-	到 "to arrive"
p^hu^{35}	p^hu^{31}	p^hou^{35}	$*p^h-$	普 "normal"
t^hu^{35}	t^hu^{31}	t^hou^{35}	$*t^h-$	土 "earth"

Another example lies in the development of ancient tones. The Dapeng dialect separates the Middle Chinese Qù tone 去聲 into the Yīn 陰 and the Yáng 陽 registers, the same as in Cantonese (§3.4.3). Hakka, on the other hand, does not distinguish the two registers of the Qù tone. The cross-dialectal comparison is illustrated in Table 22, with the numbers indicating the pitch values.

Dapeng	Hakka	Cantonese	Register
22	52	33	*Yīn* 陰
54		22	*Yáng* 陽

Table 22. The Modern Reflexes of the Middle Chinese *Qù* Tones

Compared with its sound system which presents a complex hybrid of both input dialects, the Dapeng lexicon present more similarity with Cantonese. For instance, among the approximately 100 distinctive lexical items that contain different cognates in Hakka and Cantonese, the Dapeng dialect only shares about 10% of the cognates with Hakka, while more than 80% are shared by Dapeng and Cantonese and the rest are special to Dapeng, as discussed in §4.2.6.

192

In terms of syntax, the Dapeng lexicon is slightly more similar to Cantonese than it is to Hakka. For instance, the Dapeng dialect allows the omission of the numeral or the demonstrative adjective in front of the classifier in contexts where the reference to the noun is clear enough, which is also a distinctive feature in Cantonese. Sentence (36) is an example of such omission, originally (5) in §4.3.1.

(36)　隻　　細蚊仔　　　　繼續　　　去　　搵　　隻
　　　tsik^{54}　sɐi^{22} mɐn^{42} tsɐi^{35}　kɐi^{22} tsʰɔk^{54}　hi^{22}　wɐn^{35}　tsik^{54}
　　　CL　　child　　　　　continue　　go　　find　　CL

蛤乸
kɐp^{42} na^{35}
frog

"(這/那)個孩子繼續去找(這/那)隻青蛙"
'The child continues looking for the frog'

Apart from the analyses of the hybrid nature of Dapeng, the ample, first-hand fieldwork data included in this dissertation also facilitates future research on this local dialect. In particular, since the description is written following the well-established format of the conventional "dialect report," it also enables both future synchronic and diachronic cross-dialectal comparisons with other Chinese dialects recorded in the same framework.

7.1.2 Summary of the Dapeng Formation Processes

Built upon the detailed description of the mixed nature of the Dapeng dialect, the dissertation has answered the second research question, i.e. the formation processes of the Dapeng dialect. Chapter 5 introduces and discusses Trudgill's (1986) model of "koineization." This chapter also clarifies some ambiguous use of terminology in the literature (§5.1). Most importantly, it is proposed in this chapter that koineization *per se* is not a unique linguistic process but is a combination of two separate processes: levelling and simplification. Based on this understanding, the model of "koineization" is employed to account for the formation processes of the contemporary Dapeng dialect, with particular focuses on how the two linguistic processes have shaped the Dapeng structure (§5.4).

In the process of levelling, the Dapeng dialect relied heavily on the relative salience of the linguistic features in the two contributing dialects. The strategy of choosing features is always to adopt the less salient ones and to eliminate the more marked ones between Hakka and Cantonese. For instance, the Dapeng system avoided the salient phoneme /œ/ in Cantonese altogether, in which case Dapeng tended to follow the corresponding Hakka variants. The correspondence is shown in Example (37), originally (28) in §5.4.1.

(37)

Dapeng	Hakka	Cantonese	Morpheme
hiɔ^{42}	hiɔ^{44}	hœ^{53}	靴 "boot"
k^{h}i^{22}	k^{h}i^{44}	køy^{22}	巨 "huge"
tsɐn^{22}	tsun^{52}	tʃøn^{33}	俊 "handsome"
iɔŋ^{31}	iɔŋ^{11}	jœŋ^{21}	陽 "sun"
siɔk^{42}	siɔk^1	sœk^{33}	削 "pare"

Another case is the modern reflexes of Middle Chinese 梗 Gěng final group. The Dapeng dialect evidently followed the more regular vowel distribution pattern in Cantonese rather than the more irregular one in Hakka. The patterns are shown in (38), originally (29) in §5.4.2.

(38) Modern reflexes of the Middle Chinese 梗 Gěng final group
(a)

Dapeng	Hakka	Cantonese	Morpheme
maŋ^{31}	maŋ^{11}	maŋ^{21}	盲 "blind"
pak^{42}	pak^1	pak^{33}	百 "hundred"
saŋ^{35}	sɛn^{31}	ʃaŋ^{35}	省 "province"
ts^{h}ak^{54}	ts^{h}ɛt^5	tʃak^2	宅 "house"
ts^{h}aŋ^{54}	ts^{h}aŋ^{52}	tʃɛŋ^{22}	鄭 "Zheng (surname)"
ts^{h}ak^{42}	ts^{h}ak^1	tʃ^{h}ɛk^{33}	尺 "ruler"

(b)

Dapeng	Hakka	Cantonese	Morpheme
lɐŋ^{31}	laŋ^{11}	liŋ^{21}	零 "zero"
tɐŋ^{42}	tɛn^{44}	tiŋ^{53}	丁 "labor"
sɐŋ^{42}	sən^{44}	ʃiŋ^{53}	星 "star"
sɐk^{42}	sət^1	ʃik^5	釋 "release"
kɐŋ^{42}	kin^{44}	kiŋ^{53}	京 "capital"
pɐk^{42}	pit^1	pik^5	璧 "jade"

195

In both cases of avoiding /œ/ and adopting the more regular vowel distribution pattern, the Dapeng dialect levelled out the differences between the input dialects by eliminating the more salient, marked feature between the two.

In addition to levelling, this study has also proposed simplification as another linguistic process contributing to the formation of Dapeng. Through simplification, structural complexity and irregularities were reduced in the newly formed Dapeng dialect. For instance, the somewhat redundant multiple layers of phonological forms which brought in undesirable complexity to the dialect system were lost in Dapeng, as shown in (39), originally (34) in §5.4.4.

(39) Multiple layers of pronunciation in modern Chinese dialects

Dapeng	Hakka	Cantonese	Morpheme
kʰɐn^{22}	kʰiun^{44}, kʰiun^{52}	kʰɐn^{23}, kɐn^{22}	近 "close"
tʰin^{22}	tʰɔn^{44}, tʰɔn^{52}/tʰɔn^{31}	tʰ^{23}, tyn^{22}	斷 "break (v.)"
tsʰɔ^{22}	tsʰɔ^{44}, tsʰɔ^{52}	tʃʰ^{23}, tʃɔ^{22}	坐 "sit"
tʰɐi^{54}	tʰai^{44}, tʰi^{52}	tɐi^{22}	弟 "younger brother"
tʰuŋ^{54}	tʰuŋ^{44}, tʰuŋ^{52}	tʊŋ^{22}	動 "move (v.)"

Another example is the merger of the progressive and continuous imperfective aspects in the Dapeng dialect, which are still kept distinctive in both Hakka and Cantonese. The Dapeng dialect uses the same post-verbal marker 緊 /kɐn^{35}/ as both the progressive and continuous imperfective markers. The merger is illustrated in Sentences (40) and (41), originally (23) and (24) in §4.3.7.

(40) (佢) 養到 隻 蛤嫲 … 就 放緊

(kʰi^{22}) ioŋ^{35} tau^{22} tsik^{54} kɐp^{42}na^{35}… tsʰɐu^{54} fɔŋ^{22} kɐn^{35}

(3rd SG) raise-CONT CL frog then put-CONT

在 房間

tsʰui^{54} fɔŋ^{31} kan^{42}

in room

"(他)養著一隻青蛙 … 就(把它)放在房間裡"

'(He) has a frog… and puts it in the room'

(41) 貓頭鷹 追緊 佢哋

miu^{35} tʰɐu^{31} jɐŋ^{42} tsui^{42} kɐn^{35} kʰi^{22} tʰi^{54}

Owl chase-PROG 3rd pl

"貓頭鷹在追他們"

'The owl is chasing them'

In sum, the Dapeng dialect adopted features from contributing dialects based on salience and markedness. In the process of simplification, the Dapeng dialect further adapted newly adopted features to achieve lower complexity and higher regularity.

Based on linguistic evidence, this chapter has hence argued that the Dapeng dialect was formed through the process of "koineization," which includes two more general linguistic processes: levelling and simplification. This hypothesis was further backed up by socio-historical evidence, specifically the migration history of the Dapeng area (§5.3).

7.1.3   Summary of the Dapeng Dialect Vitality

The third and last research question pertains to the vitality of the Dapeng dialect in today's peninsula community. To address this question about language vitality as well

as linguistic and social factors, this dissertation first reviewed some of the most influential frameworks that have been proposed for language vitality assessment (§6.2) and has demonstrated that the UNESCO Language Vitality and Endangerment (LVE) scale is the most suitable and most applicable one in the Chinese context (§6.3). Then this model was employed to the evaluation of the vitality of the Dapeng dialect, supported by evidence drawn from first-hand interviews, observations, and demographic data collected during fieldwork.

The results of assessment have shown that, despite the low scores of some of the factors in the UNESCO-LVE scale, which might suggest a fairly low vitality, an overview of all factors clearly suggests an overall vigorous condition. More specifically, in spite of inadequate institutional support and less developed literature/orthographies, the majority of the factors all point to a positive, vital situation of local dialect use in the Dapeng community.

This condition makes Dapeng distinctive from many other small local Chinese dialects, which are usually reported in the literature as in danger (§6.4). Through this assessment, this chapter has shown that not all small dialects in China are severely endangered. One will have to rely on evaluative frameworks in order to have a comprehensive, systematic view of the vitality issue of a dialect, which could be heavily determined by the very specific sociolinguistic ecology in the local community.

7.2 Some Limitations of This Study

As probably the first detailed study dedicated to the undocumented Dapeng dialect, this dissertation has some limitations. First, although 20 native speakers of Dapeng were interviewed across different gender and age groups (the 20's through the 80's), only a portion of those data are used in this dissertation due to the restrictions of volume and time. For instance, in the description of the Dapeng sound system, only one of the male senior speakers' pronunciation is thoroughly documented. While the traditional "dialect report" normally does not require more than one speaker, analyzing the pronunciations of more speakers from various age groups could potentially show a clearer picture as to how the Dapeng sound system has changed over recent decades, which is yet another research topic. Future studies can continue and investigate the Dapeng dialect from this perspective, as further discussed below in §7.3.

Another limitation is the relatively short length of the fieldwork. Two months perhaps were enough to collect data of the local dialect, but they may not be enough for a researcher to have a full understanding of the dialect speech community. For instance, there was no proper opportunity during the fieldwork trip to closely observe the language use of native speakers at the K-12 age, whose linguistic ability and language attitudes are both highly malleable. Relevant information was gathered from interviews with their parents' and grandparents' generation, but that is not as good as direct observation. Our understanding of the intergenerational language transmission of the Dapeng dialect could

199

have been more thorough if direct contact with and direct observation of elementary students and adolescents were available.

7.3   Directions for Future Studies

First, as mentioned above, future research can place more focus on the cross-generational change of the Dapeng dialect, for instance, how the Dapeng sound system has changed in recent decades. Dialect change is a critical issue in the Chinese setting due to the continuous promotion, for over half a century, of Putonghua and to the thriving of the mass media in recent decades.

In the case of Dapeng, as the majority of Dapeng speakers can speak both Cantonese and Putonghua (and sometimes Hakka as well) in addition to their native dialect, the situation is even more complicated. The influence of these major, powerful dialects may have impacted each generation in different ways. Therefore, the study of the pervasive multidialectalism with the emphasis on generational differences could greatly contribute to scholars' understanding of how a local dialect interacts with more powerful dialects—such as Hakka, Cantonese, and Putonghua—and how it is changed or molded accordingly.

Second, research on the Dapeng dialect can continue and expand from the peninsula community in Shenzhen to the oversea Dapeng communities, for instance, the one in New York City. It could be fairly interesting to see how the Dapeng dialect has changed with the long term contact with non-Chinese host languages, such as English. Results from the research on the oversea communities can then be compared in

juxtaposition with the results in this dissertation. The comparisons of both the Dapeng language structure and the dialect speech communities would be of great interest to the field of Chinese dialectology.

Even one more step away from the current study, future comparative studies of the Dapeng dialect could move from the traditional, primarily qualitative methods in this dissertation to a quantitative approach, namely, dialectometry. Dialectometry is the measurement of linguistic differences at all lexical, phonological, and morphosyntactic levels among speech varieties. It is a method that quantifies dialect "distances" based on a large amount of synchronic data (Cheng 1997, Heeringa 2004, Nerbonne 2009).

As a data-driven, aggregated approach, dialectometry does not require predetermined knowledge of dialects. That is, there is no need to select among individual data points or features for comparison, but rather distances among all dialect points may be calculated based on the whole data set available. Aggregating the differences and similarities of all data points and all features could show overall distances and affinities across dialects.

In the case of Dapeng, if data of both the local dialect and from the input dialects are analyzed with an aggregated approach, one could calculate and visualize dialect relationships with the help of computer software. The quantitative approach is a great complement of the conventional, primarily qualitative methodology in this dissertation,

and they should together contribute to the comparative studies of the Dapeng dialect and

the other neighboring Southern Chinese dialects.[90]

---

[90] So far this approach has been applied to the comparative dialect study in the Chinese context both at the national level (Wang 1994, Cheng 1997, Hamed 2005, Hamed and Wang 2007, Tang 2009, etc.) and at the regional level in Southern China (Chen 2012 on Shaoguan Tuhua 韶關土話, Tsui and Chen forthcoming on the Xiang-Gan-Hakka relationship, etc.).

References

Aikhenvald, Alexandra Y. and Robert M. W. Dixon (eds.). 2001. *Areal diffusion and genetic inheritance: problems in comparative linguistics*. Oxford: Oxford University Press.

Ansaldo, Umberto. 2010. Contact and Asian Varieties of English. In: Raymond Hickey (ed.) *The Handbook of Language Contact*. Malden, MA/ Oxford: Blackwell. Pages 498-517.

Auer, Peter, Frans Hinskens, and Paul Kerswill (eds.). 2005. *Dialect Change: Convergence and Divergence in European Languages*. Cambridge [u.a.]: Cambridge University Press.

Baoan County Annals Committee (寶安縣地方史志编纂委员会) (ed). 1997. *Baoan Xianzhi: Junshi Zhi* (寶安縣志) (The annals of Baoan County). Guangzhou: Guangdong Renmin Chubanshe.

Bates, Peter, Moe Chiba, Sabine Kube, and Douglas Nakashima, eds. 2009, *Learning and Knowing in Indigenous Societies Today*. UNESCO: Paris.

Beijing Daxue (北京大學中國語言文學系語言學教研室編). 2003. *Hanyu Fangyin Zihui* (漢語方音字匯) [A syllabary of Chinese dialects]. Second edition. Beijing: Wenzi Gaige Chubanshe.

Beijing Daxue (北京大學中國語言文學系語言學教研室編). 1995. *Hanyu Fangyan Cihui* (漢語方言詞匯) [A lexicon of Chinese dialects]. Beijing: Yuwen Chubanshe. Beijing: Wenzi Gaige Chubanshe.

Beijing Yuyan Daxue (北京語言大學語言研究所編). 2003. *Hanyu Fangyan Dituji Diaocha Shouce* (漢語方言地圖集調查手冊) [Handbook for the Linguistic Atlas of Chinese Dialects]. Macau: Linguistics Society of Macau.

Boersma, Paul and David Weenik. 2013. Praat: doing phonetics by computer (Version 5.3.48). http://www.praat.org. Retrieved 2 May 2013.

Britain, David and Peter Trudgill. 1999. Migration, new-dialect formation and sociolinguistic refunctionalisation: reallocation as an outcome of dialect contact. *Transactions of the Philological Society* 97. 245-256.

Bradley, David. 2006. Endangered languages of China and South-East Asia. In: Cunningham et al (eds). *Language Diversity in the Pacific: Endangerment and Survival*. Clevedon, England: Multilingual Matters. Pages 112–120.

Bradley, David. 2007. Language endangerment in China and Mainland Southeast Asia. In: Matthias Brenzinger (ed.) *Language Diversity Endangered*. Berlin: Mouton de Gruyter. Pages 278-302.

Branner, David Prager. 2000. *Problems in comparative Chinese dialectology: the classification of Miin and Hakka*. Berlin ; New York: Mouton de Gruyter.

Brenzinger, Matthias. 2007. Language endangerment throughout the world. In: Matthias Brenzinger (ed.) *Language Diversity Endangered*. Berlin: Mouton de Gruyter. Pages ix-xvi.

Brenzinger, Matthias, ed. 2007. *Language Diversity Endangered*. Berlin: Mouton de Gruyter.

Cao, Shuji (曹樹基). 1997a. *Zhongguo Yimin Shi: Ming* (中國移民史：明) [The history of migration in China: Ming Dynasty]. Vol. 4. Fuzhou: Fujian Renmin Chubanshe.

Cao, Shuji (曹樹基). 1997b. *Zhongguo Yimin Shi: Qing, Minguo* (中國移民史：清、民國) [The history of migration in China: Qing Dynasty and the Republican China]. Vol. 5. Fuzhou: Fujian Renmin Chubanshe.

Cao, Zhiyun (曹志耘). 2001. Guanyu binwei hanyu fangyan de wenti (关于濒危漢語方言问题) [On the endangered Chinese dalects]. *Language Teaching and Linguistic Studies* (語言教學與研究). (1): 8-12.

Cao, Zhiyun (曹志耘) (ed,). 2008. *Hanyu fangyan ditu ji* (漢語方言地圖集) [Linguistic atlas of Chinese dialects]. Beijing: Shangwu Yinshuguan.

Chafe, Wallace (ed.). 1980. *The Pear Film: Cognitive, Cultural and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex.

Chambers, John Kenneth, Peter Trudgill, and Natalie Schilling-Estes (eds.). 2013. *The Handbook of Language Variation and Change*. Oxford: John Wiley & Sons.

Chan, Marjorie K.M. 1980. *Zhong-shan phonology: a synchronic and diachronic analysis of a Yue (Cantonese) dialect*. M.A. thesis, University of British Columbia, Vancouver, Canada.

Chan, Marjorie K.M. 1985. *Fuzhou Phonology: A Non-Linear Analysis of Tone and Stress*. Ph.D. dissertation, University of Washington, Seattle, U.S.A.

Chao, Yuen-Ren. 1930. A system of tone-letters. *Le Maître Phonétique* 45: 24–27.

Chao, Yuen-Ren. 1934. The non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology IV*, Part 4, 363-97.

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chao, Yuen Ren. 1975. My field work on the Chinese dialects. *Computational Analysis of Asian and African Languages (CAAAL)* 2: 3-7.

Chao, Yuen Ren. 1980. A system of "tone-letters." *Fangyan* (方言) 1980.2.81-83.

Chappell, Hilary. 2001. Language contact and areal diffusion in Sinitic languages. In: Alexandra Y. Aikhenvald and Robert M. W. Dixon (eds.) *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford: Oxford University Press. Pages 328-357.

Chappell, Hilary (曹茜蕾). 2007. Hanyu fangyan de chuzhi biaoji de leixing (漢語方言的處置標記的類型) [Typology of object marking constructions: A pan-Sinitic view]. *Yuyanxue Luncong* (語言學論叢), 36: 183-209.

Chappell, Hilary (ed.). 2015a. *Diversity in Sinitic Languages*. Oxford University Press.

Chappell, Hilary. 2015b. Linguistic areas in China for differential object marking, passive, and comparative constructions. In: Hilary Chappell (ed.) *Diversity in Sinitic Languages*. Oxford University Press. Pages 13-52.

Chen, Litong. 2012. *Shaoguan Tuhua, a Local Vernacular of Northern Guangdong Province, China: A New Look from a Quantitative and Contact Linguistic Perspective.* M.A. thesis, The Ohio State University, Columbus, Ohio.

Chen, Xiaojin (陳曉錦). 1993. *Dongguan Fangyan Shuolüe* (東莞方言說略) [A brief introduction to dialects in Dongguan]. Guangzhou: Guangdong renmin chubanshe.

Cheng, Chin-Chuan. 1987. A quantification of Chinese dialect affinity. *Wang Li Memorial Volumes*, English volume, 73-100.

Cheng, Lisa Lai-Shen and Rint Sybesma. 2005. Classifiers in four varieties of Chinese. In: Guglielmo Cinque and Richard Kayne (eds.) *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press. Pages 259–292.

Coblin, W. South. 2002. Migration history and dialect development in the lower Yangtze watershed. *Bulletin of the School of Oriental and African Studies* 65.3.529-543.

Comrie, Bernard and Norval Smith. 1977. Linguistic descriptive studies: questionnaire. *Lingua* 42.1-72.

Cunningham, Denis, D. E. Ingram and Kenneth Sumbuk (eds). 2006. *Language Diversity in the Pacific: Endangerment and Survival*. Clevedon, England: Multilingual Matters.

Dede, Keith. 2007. The deep end of the feature pool: Syntactic hybridization in Chinese dialects. *Journal of Chinese Linguistics* 35.1.58-80.

Deng, Xiaohua (鄧曉華) and William S-Y Wang (王士元). 2009. *Zhongguo de yuyan ji fangyan de fenlei* (中國語言及方言的分類) [The classification of Chinese languages and dialects]. Beijing: Zhonghua Shuju.

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.

Ewing, Michael, Margaret Florey, and Romola Rassool. 2013. *Developing a Tool to Assess Language Vitality*. Presentation at 3rd International Conference on Language Documentation and Conservation (ICLDC). Retrieved January 31, 2015, from http://scholarspace.manoa.hawaii.edu/handle/10125/26068.

Fishman, Joshua A. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Vol. 76. Multilingual matters.

Fishman, Joshua A. 2000. Reversing language shift: RLS theory and practice revisited. In: Gloria E. Kindell and M. Paul Lewis (eds.), *Assessing Ethnolinguistic Vitality: Theory and Practice: Selected Papers from the Third International Language Assessment Conference*. SIL Publications in Sociolinguistics. Dallas: SIL International. Pages 1–26.

Florey, Margaret. 2005. Language shift and endangerment. In: Alexander Adelaar and Nikolaus P. Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*. Routledge Language Family Series. London: Routledge. Pages 43-64

Florey, Margaret. 2007. *Building Tools for Assessing Linguistic Diversity*. Presentation at International Conference on Austronesian Endangered Language Documentation (Providence University, Taiwan). Retrieved January 8, 2015, from http://www.rnld.org/sites/default/files/Florey%20Linguistic%20Vitality%20Test.pdf.

Florey, Margaret. 2009. Sustaining indigenous languages and indigenous knowledge: Developing community training approaches for the 21st century. In: Peter Bates, Moe Chiba, Sabine Kube, and Douglas Nakashima (eds.), *Learning and Knowing in Indigenous Societies Today*. UNESCO: Paris. Pages 25-39.

Gao, Huanian (高華年). 1980. *Guangzhou Fanguan Yanjiu* (廣州方言研究) [Research on the Guagnzhou dialect]. Hong Kong: The Commercial Press.

Ghosh Johnson, Subhadra Elka. 2006. *Mexiqueño?: Issues of Identity and Ideology in a Case Study of Dialect Contact*. Ph.D. Dissertation: University of Pittsburgh.

Guangdong Province Annals Committee (广东省地方史志编纂委员会) (ed). 1999. *Guangdong Shengzhi: Junshi Zhi* (廣東省志：軍事志) [The annals of Guangdong Province: military]. Guangzhou: Guangdong Renmin Chubanshe.

Guo, Longsheng. 2004. The relationship between Putonghua and Chinese dialects. In: Minglang Zhou (ed.), *Language Policy in the People's Republic of China: Theory and Practice Since 1949*. Boston: Kluwer Academic Publishers. Pages 45-53.

Hamed, Mahé Ben. 2005. Neighbour-nets portray the Chinese dialect continuum and the linguistic legacy of China's demic history. *Proceedings of the Royal Society of London* B, 272, 1015-22.

Hamed, Mahé Ben and Feng Wang. 2006. Stuck in the forest: Trees, networks and Chinese dialects. *Diachronica* 23, 29–60.

Hashimoto, Mantaro J. 1973. *The Hakka dialect; a linguistic study of its phonology, syntax, and lexicon*. Cambridge [England]: University Press.

Hashimoto, Mantaro J. (橋本萬太郎). Zhihong Yu (余志鴻) (trans.). 2008. *Yuyan Dili Leixingxue* (語言地理類型學) [A Geographical Typology of Languages]. Beijing: World Publishing Corporation.

Handel, Zev. 2014. Historical Phonology of Chinese. In: C.-T. James Huang, Y.-H. Audrey Li, and Andrew Simpson (eds.), *The Handbook of Chinese Linguistics*. Chichester, West Sussex, UK: Wiley-Blackwell. Pages 576-598.

He, Gengyong (何耿鏞). 1993. *Kejia Fangyan Yufa Yanjiu* (客家方言語法研究) [Studies on the Hakka grammar]. Xiamen: Xiamen University Press.

Hickey, Raymond (ed.). 2010. *The Handbook of Language Contact*. Malden, M.A.: Wiley-Blackwell.

Hinskens, Frans, Peter Auer, and Paul Kerswill. The study of dialect convergence and divergence: Conceptual and methodological considerations. In: Peter Auer, Frans Hinskens, and Paul Kerswill (eds.). 2005. *Dialect Change: Convergence and Divergence in European Languages*. Cambridge [u.a.]: Cambridge University Press. Pages 1-48.

Ho, Dah-an. 2015. Chinese dialects. In: William S.-Y. Wang and Chaofen Sun (eds.), *The Oxford handbook of Chinese linguistics*. New York, NY : Oxford University Press. Pages 149-159.

Huang, C.-T. James, Y.-H. Audrey Li, and Andrew Simpson (eds.). 2014. *The Handbook of Chinese Linguistics*. Chichester, West Sussex, UK : Wiley-Blackwell.

Huang, Han-chun. 2015. Relating Causative and Passive Bun Constructions in Hakka. *Tsing Hua Journal of Chinese Studies*.45(2) 167-200

Huang, Xiaodong (黃曉東). 2007. Hanyu Junhua gaishu (漢語軍話概述) [A Survey of Junhua in China]. *Language Teaching and Linguistic Studies* (語言教學與研究). (3): 21-27.

Hudson Kam, Carla L. and Elissa L. Newport. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1, 151-195.

Ji, Zhilong (紀志龍) (ed). 2001. *Lishi Wenhua Mingzhen: Dapeng Zhen* (歷史文化名鎮：大鵬鎮) [Dapeng: A Renowned Town of History and Culture]. Hong Kong: Tiandiren Chuban Youxian Gongsi.

Kerswill, Paul. 2002. Models of linguistic change and diffusion: new evidence from dialect levelling in British English. *Reading Working Papers in Linguistics* 6. 187-216.

Kerswill, Paul. 2010. Contact and new varieties. In: Raymond Hickey (ed.) *The Handbook of Language Contact*. Malden, M.A.: Wiley-Blackwell. Pages 230–51.

Kerswill, Paul. 2013. Koineization. In: John Kenneth Chambers and Natalie Schilling-Estes (eds.) *The Handbook of Language Variation and Change*. John Wiley & Sons. Pages 519-536.

Kerswill, Paul and Peter Trudgill. 2005. The birth of new dialects. In: Peter Auer, Frans Hinskens, and Paul Kerswill (eds.). 2005. *Dialect Change: Convergence and Divergence in European Languages*. Cambridge [u.a.]: Cambridge University Press. Pages 196–220.

Kerswill, Paul and Ann Williams 2000. Creating a new town koine: children and language change in Milton Keynes. *Language in Society* 29.1: 65–115.

Kerswill, Paul and Ann Williams. 2005. New towns and koineisation: linguistic and social correlates. *Linguistics* 43.5: 1023–48.

Kindell, Gloria Elaine and M. Paul Lewis. 2000. *Assessing Ethnolinguistic Vitality: Theory and Practice: Selected Papers from the Third International Language Assessment Conference*. Dallas: SIL International.

King, Kendall A. 2008. *Sustaining Linguistic Diversity: Endangered and Minority Languages and Language Varieties*. Washington, D.C: Georgetown University Press.

Klein, Wolfgang, Ping Li, and Hemriette Hendriks. 2000. Aspect and assertion in Mandarin Chinese. *Natural Language & Linguistic Theory* 18 (4): 723-770.

Krauss, Michael. 1992. The world's languages in crisis. *Language* 68(1): 4–10.

Krauss, Michael. 2007. Classification and terminology for degrees of language endangerment. In: Matthias Brenzinger (ed.) *Language Diversity Endangered*. Berlin: Mouton de Gruyter. Pages 1-8.

Kuo, Yun-Hsuan. 2005. *New Dialect Formation: the Case of Taiwanese Mandarin*. Ph.D. dissertation, University of Essex.

Kurpaska, Maria. 2010. *Chinese Language(s): a Look through the Prism of the Great Dictionary of Modern Chinese Dialects*. Berlin: De Gruyter Mouton.

Landweer, M. Lynn. 2012. Methods of language endangerment research: a perspective from Melanesia. *International Journal of the Sociology of Language*. 2012(214): 153-178.

LaPolla, Randy J. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family. In: Alexandra Y. Aikhenvald and Robert M. W. Dixon (eds.) *Areal diffusion and genetic inheritance: problems in comparative linguistics*. Oxford: Oxford University Press. Pages 225-254.

LaPolla, R. J. (2010). Language contact and language change in the history of the sinitic languages. *Procedia-Social and Behavioral Sciences*, 2(5), 6858-6868.

Legère, Karsten. 2007. Vidunda (G38) as an endangered language?. In: Doris L. Payne and Jaime Peña (eds.), *Selected Proceedings of the 37th Annual Conference on African Linguistics*. Somerville, MA: Cascadilla Proceedings Project. Pages 43-54

Lewis, M. Paul. 2006. Towards a categorization of endangerment of the world's languages. *SIL Electronic Working Papers.* 2006-002. Retrievd January 21, 2015, from http://www-01.sil.org/silewp/2006/silewp2006-002.pdf.

Lewis, M. Paul. 2008. Evaluating endangerment: Proposed metadata and implementation. In: Kendall A. King, Natalie Schilling-Estes, Lyn Fogle, Jia Lou and Barbara Soukup (eds.), *Sustaining Linguistic Diversity: Endangered and Minority Languages and Language Varieties*. Washington, DC: Georgetown University Press. Pages 35–49.

Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com/16.

Lewis, M. Paul. 2011. *The Sustainable Use Model for Language Development*. Presentation at the American Association of Applied Linguistics, Chicago, IL, 26-29 March 2011. Retrieved February 2, 2015, from: http://www.sil.org/~simonsg/presentation/Applying%20the%20SUM.pdf.

Lewis, M. Paul and Gary F. Simons. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 55(2): 103–120.

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2014. *Ethnologue: Languages of the World*, Seventeenth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Li, Charles N. and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*, University of California Press, Berkeley.

Li, David C.S. 2015. Lingua francas in Greater China. In: William S.-Y. Wang and Chaofen Sun (eds.), *The Oxford handbook of Chinese linguistics*. New York, NY : Oxford University Press. Pages 590-600.

Li, Dongxiang (李冬香) and Zhuang Chusheng (莊初升). 2009. *Shaoguan Tuhua diaocha yanjiu* (韶關土話調查研究) [Investigation and research on Shaoguan Tuhua]. Guangzhou: Ji'nan University Press.

Li, Fang-kuei. 1973. Languages and dialects of China. *Journal of Chinese Linguistics* 1.1.1-13. (Originally published in The Chinese Yearbook (1937), Pages 59-65. Shanghai: Shangwu Yinshuguan.)

Li, Rong (李榮). 1989a. Zhongguode yuyan he fangyan (中國的語言和方言) [Languages and dialects in China]. *Fangyan* (方言) 1989.3.161-167.

Li, Rong (李榮). 1989b. Hanyu fangyande fenqu (漢語方言的分區) [The classification of the Chinese dialects]. *Fangyan* (方言) 1989.4.241-259.

Li, Rulong (李如龍). 2001. *Hanyu Fangyan de Bijiao Yanjiu* (漢語方言的比較研究) [A comparative study of Chinese dialects]. Beijing: Shangwu Yinshuguan.

Li, Rulong (李如龍) (ed.), *Hanyu Fangyan Tezhengci Yanjiu* (漢語方言特征詞研究) [Studies on distinctive words in Chinese dialects]. Xiamen: Xiamen Daxue Chubanshe.

Li, Rulong (李如龍). 2007. *Hanyu Fangyanxue* (漢語方言學) [Chinese dialectology]. Beijing: Gaodeng Jiaoyu Chubanshe.

Li, Rulong (李如龍) and Zhuang Shuangqing (張雙慶). 1992. *Ke Gan Fangyan Diaocha Baogao* (客贛方言調查報告) [A report on a survey of the Hakka and Gan dialects]. Xiamen: Xiamen University Press.

Li, Xiaofan (李小凡) and Mengbing Xiang (項夢冰). 2009. *Hanyu Fangyanxue Jichu Jiaocheng* (漢語方言學基礎教程) [Course in Chinese dialectology]. Beijing: Beijing Daxue Chubanshe.

Li, Yuming (李宇明) and Qianrui Chen (陳前瑞). 2005. Beijinghua gei zi beidongju de diwei ji qi lishi fazhan (北京話"給"字被動句的地位及其歷史發展) [The status and development of passive "gei" structures in Peking Dialect]. *Fangyan,* 4.1-9.

Liu, Danqing (劉丹青). 2008. *Yufa Diaocha Yanjiu Shouce* (語法調查研究手冊) [Handbook of Syntactic Research]. Shanghai: Shanghai Jiaoyu Chubanshe.

Liu, Jian (劉堅) (ed.). 1998. *Ershi Shiji de Zhonguo Yuyanxue* (二十世紀的中國語言學) [Linguistics in 20th century China]. Beijing: Beijing daxue chubanshe

Malmqvist, Göran. 1994. Chinese Linguistics. In: Giulio Lepschy (ed.), *History of Linguistics. Volume 1. The Eastern Traditions of Linguistics*. London and New York: Longman. Pages 1-24.

Matsumura, Kazuto, ed. 1998. *Studies in Endangered Languages: Papers from the International Symposium on Endangered Languages, Tokyo, November 18-20, 1995*. Tokyo: Hituzi Syobo.

Matthews, Stephen. 1996. Ditaxia and hybridization in Chinese dialect grammar. *Proceedings of the Fourth Conference on Pan-Asiatic Linguistics*, Bangkok.1274-1283.

Matthews, Stephen. 1999. Y. R. Chao and Universal Chinese Grammar. In: David Cram, Andrew Robert Linn and Elke Nowak (eds.), *History of Linguistics 1996*. Amsterdam: John Benjamins. Pages 217-224.

Matthews, Stephen. 2010. Language Contact and Chinese. In: Raymond Hickey (ed.) *The Handbook of Language Contact*. Malden, MA/ Oxford: Blackwell. Pages 757-769.

Matthews, Stephen and Virginia Yip. 1994. *Cantonese: A Comprehensive Grammar*. London: Routledge.

Mayer, Mercer. 2003. *Frog, where are you?* New York: Dial Books for Young Readers.

McConvell, Patrick and Nicholas Thieberger. 2001. *State of Indigenous Language in Australia –2001*. Australia State of the Environmental Technical Paper Series (Natural and Cultural Heritage), Series 2, Department of the Environment and Heritage, Canberra. Retrieved February 1, 2015, from: http://www.environment.gov.au/system/files/pages/30bedb2e-29d7-4733-b45d-835b1bf2bef3/files/indigenous-languages.pdf.

Mei, Tsu-Lin. (梅祖麟). 1995. Fangyan benzi yanjiu de liangzhong fangfa (方言本字研究的兩種方法) [Two approaches in the study of original characters of dialects]. In: Tsu-Lin Mei, *Wuyu he Minyu de Bijiao Yanjiu* (吳語和閩語的比較研究) [A comparative study of the Wu and Min dialects]. Shanghai: Shanghai Jiaoyu Chubanshe. Pages 1-12.

Mihas, Elena, Bernard C. Perley, Gabriel Rei Doval, Kathleen Wheatley, and Michael Noonan. 2013. *Responses to Language Endangerment: in Honor of Mickey Noonan: New Directions in Language Documentation and Language Revitalization*. Studies in Language Companion Series 142. Amsterdam: John Benjamins.

Minasyan, Anahit, 2011. *UNESCO's Language Vitality and Endangerment Methodological Guideline: Review of Application and Feedback since 2003*. Prepared by UNESCO's Culture Sector for expert meeting "Towards UNESCO guidelines on Language Policies: a Tool for Language Assessment and Planning" (30 May – 1 June 2011). Retrieved February 4, 2015, from: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/unesco_language_vitaly_and_endangerment_methodological_guideline.pdf.

National Taiwan University Department of Chinese Literature, Institute of History and Philology, Academia Sinica, and Information Science Institute. The Xiaoxuetang Database of Chinese Characters (小學堂文字學資料庫). http://xiaoxue.iis.sinica.edu.tw/.

Norman, Jerry. 1988. *Chinese.* Cambridge, UK: Cambridge University Press.

Norman, Jerry. 2003. The Chinese dialects: Phonology. In: Graham Thurgood and Randy J. LaPolla (eds.), *The Sino-Tibetan Languages*. Routledge: London and New York. Pages 72-83.

Norman, Jerry and W. South Coblin. 1995. A New Approach to Chinese Historical Linguistics. *Journal of the American Oriental Society*, Vol. 115, No. 4.

Obiero, Ogone John. 2010. From assessing language endangerment or vitality to creating and evaluating language revitalization programmes. *Nordic Journal of African Studies* 19(4): 201-226.

Pan, Wuyun and Hongming Zhang. 2015. Middle Chinese phonology and Qieyun. In: William S.-Y. Wang and Chaofen Sun (eds.), *The Oxford handbook of Chinese linguistics*. New York, NY: Oxford University Press. Pages 80-90.

Peng, Xiaochuan (彭小川). 1992. Yueyu yunshu "Fenyun Cuoyao" ji qi yunmu xitong (粵語韻書《分韻撮要》及其韻母系統) [The Cantonese rhyme book "Fenyun Cuoyao" and its final system]. *Journal of Ji'nan University for Nationalities (Philosophical and Social Science Edition)* (暨南學報 [社會科學版]). Vol. 4 153-159.

Pilgrim, Sarah and Jules N. Pretty, eds. 2010. *Nature and Culture: Rebuilding Lost Connections*. London: Earthscan.

Pulleyblank, Edwin G. 1984. *Middle Chinese: A Study in Historical Phonology*. Vancouver: U. of British Columbia Press.

Pulleyblank, Edwin G. 1998. Qieyun and Yunjing: The essential foundation for Chinese historical linguistics. *Journal of the American Oriental Society* 118.2.200-216.

Qiu, Xueqiang (丘學強). 2005. *Junhua Yanjiu* (軍話研究) [A study on the Army's Speech]. Beijing: Zhongguo shehui kexue chubanshe.

Ramsey, S. Robert. 1987. *The Languages of China*. Princeton: Princeton U. Press.

Sagart, Laurent. 2002. Gan, Hakka and the Formation of Chinese Dialects. In: Dah-an Ho (ed.) *Dialect Variations in Chinese*. Taipei: Academia Sinica. Pages 129-54.

Schmidt, Annette. 1990. *The Loss of Australia's Aboriginal Language Heritage*. Canberra: Aboriginal Studies Press.

Siegel, Jeff. 1985. Koines and koineization. *Language in Society* 14: 357–78.

Siegel, Jeff. 2001. Koine formation and creole genesis. In: Norval Smith and Tonjes Veenstra (eds.) *Creolization and contact*. Amsterdam: John Benjamins. Pages 175-198.

Simons, Gary F., and M. Paul Lewis. 2013. The world's languages in crisis: A 20-year update. In: Elena Mihas, Bernard Perley, Gabriel Rei-Doval and Kathleen Wheatley (eds.), *Responses to Language Endangerment: in Honor of Mickey Noonan: New Directions in Language Documentation and Language Revitalization.* Studies in Language Companion Series 142. Amsterdam: John Benjamins. Pages 3–19.

Shenzhen Bowuguan (深圳博物館) (ed). 1997. *Shenzhen Gudai Jianshi* (深圳古代簡史) [A brief history of ancient Shenzhen]. Beijing: Wenwu Chubanshe.

Smith, Norval and Tonjes Veenstra (eds.). 2001. *Creolization and contact*. Amsterdam: John Benjamins.

Song, Tsai-Shien (宋彩仙). 2008. *Kejiahua Ti Biaoji de Yanjiu* (客家話體標記的研究) [The research of Hakka aspect]. M.A. thesis, National Central University.

Spence, Justin David. 2013. *Language Change, Contact, and Koineization in Pacific Coast Athabaskan*. Ph.D. dissertation, University of California, Berkeley.

Sun, Chaofen. 2006. *Chinese: A Linguistic Introduction*. Cambridge, UK: Cambridge University Press.

Sun, Deping (孫德平). 2012. Keyinneihua: Jianghan Youtianhua de xingcheng (柯因內化：江汉油田话的形成) [Koineization: the formation of Jianghan Oilfield Dialect]. *Studies in Language and Linguistics* (語言研究).Vol. 32 No. 4, 110-116.

Sun, Hongkai (孫宏開). 2001. Guanyu binwei yuyan de wenti (关于濒危语言问题) [On the dndangered languages in China]. *Language Teaching and Linguistic Studies* (語言教學與研究). (1): 1-7.

Sun, Hongkai (孫宏開), Zengyi Hu (胡增益), and Xing Huang (黃行) (eds.). 2007. *Zhongguo de Yuyan* (中國的語言) [The languages of China]. Beijing: Shangwu Yinshuguan.

Tan, Yuanheng (譚元亨) (ed). 2010. *Guangdong Hejia Shi* (廣東客家史) [The Hakka history in Guangdong]. Guangzhou: Guangdong Renmin Chubanshe.

Tang, Chaoju. 2009. *Mutual intelligibility of Chinese dialects: an experimental approach*. Leiden: Universiteit Leiden dissertation.

214

Tang, Sze-Wing and Siu-Pong Cheng. 2014. Aspects of Cantonese Grammar. In: C.-T. James Huang, Y.-H. Audrey Li, and Andrew Simpson (eds.), *The Handbook of Chinese Linguistics*. Chichester, West Sussex, UK: Wiley-Blackwell. Pages 601-628.

Thomason, Sarah G. and Terrence Kaufman. 1988. *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.

Thurgood,Graham and Randy J. Lapolla (eds.). *The Sino-Tibetan Languages*. Routledge: London and New York.

Ting, Pang-hsin (丁邦新). 2003. Hanyu yinyun shi shang youdai jiejue de wenti (漢語音韻史上有待解決的問題) [Unsolved problems in Chinese historical phonology]. In: Dah-an Ho (何大安) (ed.): *Gu Jin Tong Sai* (古今通塞: 漢語的歷史與發展) [The prevailing and the obsolete of past and present: The history and development of Chinese]. Taipei: Zhongyang yanjiuyuan lishi yuyan yanjiusuo. Pages 1-21.

Trudgill, Peter 1986. *Dialect in Contact*. Oxford: Blackwell.

Trudgill, Peter, Elizabeth Gordon, Gillian Lewis, and Margaret Maclagan 2000. Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics* 36, 299–318.

Tsui, Tsz-Him and Litong Chen. Forthcoming. A new quantitative analysis of the Xiang-Gan-Hakka relationship. *Proceedings of the 25th North American Conference on Chinese Linguistics (NACCL-25)*.

Tsunoda, Tasaku. 2005. *Language Endangerment and Language Revitalization*. Berlin: Mouton de Gruyter.

Tuten, Donald N. 2003. *Koineization in Medieval Spanish*. Berlin: Mouton de Gruyter.

UNESCO Ad Hoc Expert Group on Endangered Languages. 2003. *Language vitality and endangerment*. Paris, UNESCO Ad Hoc Expert Group Meeting on Endangered Languages. Accessed November 4, 2014, at http://www.unesco.org/culture/ich/doc/src/00120-EN.pdf.

Wang, Futang (王福堂). 1998. *Ershi Shiji de Zhonguo Fangyanxue* (二十世紀的中國方言學) [Dialectology in 20th century China] In: Jian Liu (劉堅) (ed.), *Ershi Shiji de Zhonguo Yuyanxue* (二十世紀的中國語言學) [Linguistics in 20th century China]. Beijing: Beijing daxue chubanshe. Pages 507-536.

Wang, Li (王力), 1980. *Hanyu Shi Gao: Shang* (漢語史稿: 上) [Monographs on the history of the Chinese language, Volume I]. Beijing: Zhonghua shuju.

Wang, Li (王力), 1987. *Wang Li Wenji Di Shi Juan: Hanyu Yuyin Shi* (王力文集 第十卷: 漢語語音史) [Collected works of Wang Li, Volume X, A history of Chinese Phonology]. Ji'nan: Shandong Jiaoyu Chubanshe.

Wang, William S-Y. 1989. Language in China: A chapter in the history of linguistics. *Journal of Chinese Linguistics* 17.2.183-222.

Wang, William S. Y. 1994. Glottochronology, Lexicostatistics, and Other Numerical Methods. *The Encyclopedia of Language and Linguistics*, Vol. 3. New York: Pergamon Press.

Wang, William S.-Y. and Chaofen Sun (eds.). 2015. *The Oxford handbook of Chinese linguistics*. New York, NY : Oxford University Press.

Wang, William S-Y. and R. E. Asher. 1995. Chinese linguistic tradition. In: E. F. K. Koerner and R. E. Asher (eds.): *Concise History of the Language Sciences: From the Sumerians to the Cognitivists*. New York: Pergamon. Pages 41-45.

Wang, Yuquan (王毓銓). 1965. *Mingdai de Juntun* (明代的軍屯) [Military colonies in the Ming dynasty]. Beijing: Zhonghua shuju.

Wei, Shuguan (韋樹關). 2006. Zhongguo binwei yuyan yanjiu xin jinzhan (中国濒危语言研究新进展) [Progresses of studies on endangered Chinese languages]. *Journal of Guangxi University for Nationalities (Philosophical and Social Science Edition)* (廣西民族大學學報 [社會科學版]). Vol. 28 (5): 2-5.

Wen, Changyan (溫昌衍). 2002. Kejia fangyan de tezhengci (客家方言的特征詞) [Distinctive words in the Hakka dialect]. In: Rulong Li (ed.), *Hanyu Fangyan Tezhengci Yanjiu* (漢語方言特征詞研究) [Studies on distinctive words in Chinese dialects]. Xiamen: Xiamen Daxue Chubanshe. Pages 208-277.

Williams, Ann and Paul Kerswill. 1999. Dialect levelling: change and continuity in Milton Keynes, Reading and Hull. In: Paul Foulkes and Gerry Docherty (eds.), *Urban Voices: Accent Studies in the British Isles*. London: Arnold. Pages 141–62.

Winford, Donald. 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell.

Winford, Donald. 2005. Contact-induced changes: classification and processes. *Diachronica* 22(2), 373-427.

Winford, Donald. 2007. Some issues in the study of language contact. *Journal of Language Contact (JLC)* Vol. 1

Wu, Yicheng and Adams Bodomo. 2009. Classifiers ≠ determiners. *Linguistic Inquiry* 40: 487-503.

Wu, Yonghuan (吳永煥). 2008. Hanyu fangyan wenhua yichan baohu de yiyi yu duice (漢語方言文化遺產保護的意義與對策) [On significance and strategies of protecting the cultural heritage of Chinese dialects]. *Journal of Renmin University of Chia)* (中國人民大學學報). (4): 39-43.

Wurm, Stephen A, et al. (eds). 1987. Chinese dialects in the Guangdong Province (廣東省的漢語方言). In: *Language Atlas of China* (中國語言地圖集). Hong Kong: Longman (Far East) Ltd.

Wurm, Stephen A. 1998. Method of language maintenance and revival, with selected cases of language endangerment in the world. In: Kazuto Matsumura (ed.), *Studies in Endangered Languages: Papers from the International Symposium on Endangered Languages, Tokyo, November 18-20, 1995*. Tokyo: Hituzi Syobo. Pages 191-212.

Wurm, Stephen A. 2007. Threatened languages in the Western Pacific area from Taiwan to, and including, Papua New Guinea. In: Matthias Brenzinger (ed.) *Language Diversity Endangered*. Berlin: Mouton de Gruyter. Pages 374-390.

Xiong Zhenghui (熊正輝) and Zhang Zhenxing (張振興). 2008. Hanyu fangyan de fenqu (漢語方言的分區) [Classification / Distribution of Chinese Dialects]. *Fangyan*, 2.97-108.

Xiong Zhenghui (熊正輝), Zhang Zhenxing (張振興), and Huang Xing (黃行). 2008. Zhongguo de yuyan (中國的語言) [Languages of China]. *Fangyan,* 3.193-203.

Yamamoto, Akira, Matthias Brenzinger and María E. Villalón. 2008. A place for all languages: on language vitality and revitalization. *Museum International*, 60: 60–70.

Yang, Hsiiu-fang. (楊秀芳). 2000. Fangyan benzi yanjiu de guannian yu fangfa (方言本字研究的觀念與方法) [Concepts and methods in the study of original characters of dialects]. *Hanxue Yanjie* (漢學研究) [Chinese Studies]. Special edition 18: 111-146.

Yang, Wenbo (楊文波). 2013. *Jiangxi Shangrao Tieluhua Diaocha Yanjiu* (江西上饒鐵路話調查研究) [A study on the Railway Dialect in Shangrao, Jiangxi Province]. Ph.D. dissertation: Fudan University.

Yang, Yaolin (楊耀林) and Chongyue Huang (黃崇岳). 2001. Dapeng cheng yu yapian zhanzheng (大鵬城與鴉片戰爭) [The Dapeng fortress and the Opium War]. In: Zhilong Ji (紀志龍) (ed). *Lishi Wenhua Mingzhen: Dapeng Zhen* (歷史文化名鎮：大鵬鎮) [Dapeng: A Renowned Town of History and Culture]. Hong Kong: Tiandiren Chuban Youxian Gongsi. Pages 153-167.

You, Rujie (游汝杰). 1991. *Hanyu Fangyanxue Daolun* (漢語方言學導論) [Introduction to Chinese dialectology]. Shanghai: Shanghai Jiaoyu Chubanshe

You, Rujie (游汝杰). 1993. Heilongjiang sheng de Zhanren he Zhanhua shulüe (黑龍江省的站人和站話述略) [A survey of Zhanhua and its speakers in Heilongjiang Province]. *Fangyan* (方言). (2): 142 -147.

Yuan, Jiahua (袁家驊). 2001. *Hanyu Fangyan Gaiyao* (漢語方言概要) [An outline of Chinese dialects]. Beijing: Yuwen Chubanshe.

Yue-Hashimoto, Oi-Kan. 1972. *Studies in Yue Dialects 1: Phonology of Cantonese*. Cambridge: Cambridge University Press.

Yue-Hashimoto, Anne Oi-Kan. 1993. *Comparative Chinese Dialectal Grammar: Handbook for Investigators*. Paris: Centre de Recherches Linguistiques sur l'Asie Orientale.

Zeng, Xiangwei (曾祥委). 2011. Qingchu Xin'an qianhai fujie hou de Kejia yimin chao (清初新安遷海復界后的客家移民潮) [The wave of Hakka migration to Xin'an after the abolition of the Great Evacuation in the early Qing dynasty]. *The Periodical of Hakka Research* (客家研究輯刊). (1): 70-78.

Zhan, Bohui (詹伯慧). 2002. *Guangdong Yue Fangyan Gaiyao* (廣東粵方言概要) [An outline of Yue dialects in Guangdong]. Guangzhou: Ji'nan University Press.

Zhan, Bohui (詹伯慧), Rulong Li (李如龍), Jiajiao Huang (黃家教), and Baohua Xu (許寶華). 1991. *Hanyu Fangyan Ji Fangyan Diaocha* (漢語方言及方言調查) [Chinese dialects and the investigation of dialects]. Wuhan: Hubei Jiaoyu Chubanshe.

Zhang, Guangyu (張光宇). 1990. *Ch'ieh Yün Yü Fang Yen* (切韵與方言) [Qieyun and dialects]. T'ai-pei shih: T'ai-wan shang wu yin shu kuan.

Zhang, Huiying (張惠英) (trans.) 1995. *Hanyu Gaishuo* (漢語概說) [Chinese]. Beijing: Yuwen Chubanshe.

Zhang, Shuangqing (張雙慶). 2002. Yueyu de tezhengci (粵語的特征詞) [Distinctive words in Cantonese]. In: Rulong Li (ed.), *Hanyu Fangyan Tezhengci Yanjiu* (漢語方言特征詞研究) [Studies on distinctive words in Chinese dialects]. Xiamen: Xiamen Daxue Chubanshe. Pages 390-414.

Zhang, Yibing (張一兵). 2006. *Shenzhen Jiuzhi Sanzhong* (深圳舊志三種) [Three Annals of Shenzhen]. Shenzhen: Haitian Chubanshe.

Zhang, Yibing (張一兵). 2007. Shenzhen fangzhi kaolüe (深圳方志考略) [Research on and summaries of Shenzhen annals]. In: Shenzhen Museum (ed) *Shenzhen Wenbo Luncong* (深圳文博論叢) [Shenzhen Culture and Museum Forum]. Beijing: Wenwu Chubanshe. Pages 113-143.

Zhao, Tong (趙彤). 2007. "Fenyun Cuoyao" de shengmu wenti (《分韵撮要》的声母问题) [On the initials of Fenyun Cuoyao]. *Linguistic Research* (語文研究) (1): 57-60.

Zhongguo Shehui Kexueyuan Yuyan Yanjiusuo (中國社會科學院語言研究所). 1981. *Fangyan Diaocha Zibiao* (方言調查字表) [Dialect survey character list]. Beijing: Shangwu Yinshuguan.

Zhou, Zhenhe (周小兵). 1997. Guangzhouhua liangci de dingzhi gongneng (廣州話量詞的定指功能) [The referential qualities of Cantonese classifiers]. *Fangyan* (方言) 1: 42-47.

Zhou, Zhenhe (周振鶴) and Rujie You (游汝杰). 2006. *Fangyan yu Zhongguo Wenhua* (方言与中国文化) [Dialect and Chinese Culture]. Second edition. Shanghai: Shanghai Renmin Chubanshe.

Zhu, Yongzhong, Üjiyediin Chuluu, Keith Slater and Kevin Stuart. Gangou Chinese Dialect. 1997. A Comparative Study of a Strongly Altaicized Chinese Dialect and Its Mongolic Neighbor. *Anthropos*. Bd. 92, H. 4./6.: 433-450.

Appendix A: Interview Questions during Fieldwork

Part I. Background information

Age 年龄: _____ Gender 性别: _____ Occupation 职业: _____
Education 教育程度: □-None 无 □-Elementary 1-3 小学 1-3 年级 □-Elementary 4-6 小学 4-6 年级
□-Middle school 初中 □-High school 高中 □-Undergraduate or higher 大学及更高
Places lived before (other than the local town) 除本镇外居住过的地区:

_____

Language/dialect-1 语言/方言一: _____ Able to speak it since 何时起开始说这种话: _____
Ability (0-N/A; 1-lowest; 9-highest; same criteria below.) 语言能力 (0-不适用; 1-最低; 9-最高; 下同):
Listening 听____ Speaking 说____ Reading 读____ Writing 写____ thinking 想___ Calculating 算____
Used where, with whom, how often 使用频率、场合及对象:

_____   _____
_____   _____
_____   _____
_____   _____

Language/dialect-2 语言/方言二: _____ Able to speak it since 何时起开始说这种话: _____
Ability (0-N/A; 1-lowest; 9-highest; same criteria below.) 语言能力 (0-本问不适用; 1-最低; 9-最高; 下同):
Listening 听____ Speaking 说____ Reading 读____ Writing 写____ thinking 想___ Calculating 算____
Used where, with whom, how often 使用频率、场合及对象:

_____   _____
_____   _____
_____   _____
_____   _____

Language/dialect-3 语言/方言三: _____ Able to speak it since 何时起开始说这种话: _____
Listening 听____ Speaking 说____ Reading 读____ Writing 写____ thinking 想___ Calculating 算____
Used where, with whom, how often 使用频率、场合及对象:

_____   _____
_____   _____
_____   _____
_____   _____

Language/dialect-4 语言/方言四: _____ Able to speak it since 何时起开始说这种话: _____
Listening 听____ Speaking 说____ Reading 读____ Writing 写____ thinking 想___ Calculating 算____
Used where, with whom, how often 使用频率、场合及对象:

_____   _____
_____   _____
_____   _____
_____   _____

Most fluent language/dialect 最精通的语言/方言: _____

Language(s)/dialect(s) used at home with parents 家中与父母说话时使用: _____

Language(s)/dialect(s) used at home with siblings 家中与兄弟姐妹说话时使用: _____

Language(s)/dialect(s) used at home with spouse 家中与配偶说话时使用: _____

Language(s)/dialect(s) used at home with children 家中与子女说话时使用: _____

Language(s)/dialect(s) used at work place 工作时使用: _____

Language(s)/dialect(s) used at school in class 学校里上课时使用: _____

Language(s)/dialect(s) used at school after class 学校里课外使用: _____

Language(s)/dialect(s) used in everyday life around the local neighborhood (e.g. with friends, with neighbors, at market, ect.) 在本地社区中（如与朋友、邻居说话，或在市场等）日常使用:

_____

Notes 备注:

_____
_____
_____
_____

Part II. Reading tasks

1. Please read the first 460 characters from the *Dialect survey character list* (Zhongguo Shehui Kexueyuan, 1981, x-xii,). 请朗读《方言调查字表》第 x 至第 xii 页的汉字。

2. Please read the following part of the *Three Character Classic*. 请诵读《三字经》节选。
   人之初，性本善。性相近，习相远。苟不教，性乃迁。教之道，贵以专。昔孟母，择邻处。
   子不学，断机杼。窦燕山，有义方。教五子，名俱扬。养不教，父之过。教不严，师之惰。
   子不学，非所宜。幼不学，老何为。玉不琢，不成器。人不学，不知义。为人子，方少时。
   亲师友，习礼仪。香九龄，能温席。孝于亲，所当执。融四岁，能让梨。弟于长，宜先知。
   首孝弟，次见闻。知某数，识某文。一而十，十而百。百而千，千而万。三才者，天地人。
   三光者，日月星。三纲者，君臣义。父子亲，夫妇顺。曰春夏，曰秋冬。此四时，运不穷。
   曰南北，曰西东。此四方，应乎中。曰水火，木金土。此五行，本乎数。曰仁义，礼智信。
   此五常，不容紊。稻粱菽，麦黍稷。此六谷，人所食。马牛羊，鸡犬豕。此六畜，人所饲。
   曰喜怒，曰哀惧。爱恶欲，七情具。口土革，木石金。与丝竹，乃八音。高曾祖，父而身。
   身而子，子而孙。自子孙，至元曾。乃九族，而之伦。父子恩，夫妇从。兄则友，弟则恭。
   长幼序，友与朋。君则敬，臣则忠。此十义，人所同…犬守夜，鸡司晨。苟不学，曷为人。
   蚕吐丝，蜂酿蜜。人不学，不如物。幼而学，壮而行。上致君，下泽民。扬名声，显父母。
   光于前，裕于后。人遗子，金满嬴。我教子，惟一经。勤有功，戏无益。戒之哉，宜勉力。

*   This classic text was written in the 13th century. It consists of a series of couplets of three characters. The complete text, using less than 1200 characters, enumerates all of the salient features of the Confucian tradition. The meaning of the first 84 characters, for example, is "<u>Men at their birth are naturally good. Their natures are much the same; their habits become widely different. If foolishly there is no teaching, the nature will deteriorate. The right way in teaching, is to attach the utmost importance to thoroughness. Of old, the mother of Mencius chose a neighborhood; and when her child would not learn, she broke the shuttle from the loom. Dou Yanshan had the right method. He taught five sons, each of whom raised the family reputation. To feed without teaching is the father's fault. To teach without severity is the teacher's laziness. If the child does not learn, this is not as it should be. If he does not learn while young, what will he be when old? If jade is not polished, it cannot become a thing of use. If a man does not learn, he cannot know his duty towards his neighbor.</u>" The translation of the rest of the *Three Character Classic* in the current task is omitted but the basic meaning is consistent with the previously translated part.

3. Please read the following short essay. 请朗读以下短文。
当汽车在望不到边际的高原上奔驰，扑入你的视野的，是黄绿错综的一条大毡子；黄的，那是土，未开垦的处女土，几十万年前由伟大的自然力所堆积成功的黄土高原的外壳；绿的呢，是人类战胜自然的结果，是麦田，和风吹送，翻起了一轮一轮的绿波——这时你会真心佩服昔人所造的两个字"麦浪"，若不是妙手偶得，便确是经过锤炼的语言的精华；黄与绿主宰着，无边无垠，坦荡如砥，这时如果不是宛若并肩的远山的连峰提醒了你（这些山峰凭你的肉眼来判断，就知道是在你脚底下的），你会忘记了汽车是在高原上行驶。这时你涌起来的感想也许是"雄壮"，也许是"伟大"，诸如此类的形容词；然而同时你的眼睛也许觉得有点倦怠，你对当前的"雄壮"或"伟大"闭了眼，而另一种味儿在你的心头潜滋暗长了——"单调"！（节选自茅盾《白杨礼赞》）
*   **Translation:** <u>When you travel by car through Northwest China's boundless plateau, all you see before you is something like a huge yellow-and-green felt blanket. Yellow is the soil—the uncultivated virgin soil. It is the outer covering of the loess plateau accumulated by Mother Nature several hundred thousand years ago. Green are the wheat fields signifying man's triumph over nature. They become a sea of rolling green waves whenever there is a soft breeze. One is here reminded of Chinese expression *mai lang* meaning "rippling wheat" and cannot help admiring our forefathers' ingenuity in coining such a happy phrase. It must have been either the brainwave of a clever scholar, or a linguistic gem sanctioned by long usage. The boundless highland, with dominant yellow and green, is flat like a whetstone. Were it not for distant mountain peaks standing side by side (which, as your naked eyes tell you, are bellow where you stand), you would probably forget that you are on the highland. The sight of the scene will probably call up inside you a string of epithets like "spectacular" or "grand". Meanwhile, however, your eyes may become weary of watching the same panorama, so much so that you are oblivious of its being spectacular or grand. And you may feel monotony coming on.</u> (From Mao Dun's *Tribute to the White Poplar*)

4. (For one or two of the elder literates only.) Please finish the rest of the *Dialect survey character list* (Zhongguo Shehui Kexueyuan, 1981). 请完成《方言调查字表》剩下的部分。

Part III. Speaking tasks

1. Please name the objects in the pictures. 请说出每张图片上画了什么。

2. Please read the 32-page wordless picture book (*Frog, Where are You?* Mayer 2003), and then restate the story in the Dapeng Dialect. 请看 32 页的图画书《青蛙，你在哪儿？》并用大鹏话复述故事情节。

3. Please watch the 6-minute film (*The Pear Film*, Chafe 1980, URL: http://pearstories.org/pears_video.htm), and then restate the story in the Dapeng Dialect. 请观看 6 分钟的小电影《梨的故事》并用大鹏话复述故事情节。

4. Please tell the story *The North Wind and the Sun* in the Dapeng Dialect. 请用大鹏话讲讲《北风和太阳》的故事。

Part IV. Spontaneous speech

1. Please briefly tell the history of the Dapeng Fortress. 请简单介绍一下大鹏所城的历史。

2. Please briefly describe how you celebrated the past Spring Festival. 请简述一下今年过年的情况。

3. General conversation
   * Description: Topics vary but will mainly on activities and everyday life in the local community. A local assistant will be hired to lead the conversation with several other local people from his or her network. At least one of the investigators will be present throughout the whole conversation to ensure no private or personal matters are discussed.)

Appendix B: Speaking Task Transcription

Frog, Where Are You? (Male, 67 years old)

呢　　　隻　　　細蚊仔　　　　　　養到　　　　　一隻　　　　蝲仔，
ni⁵⁴　　tsik⁵⁴　sei²² men⁴² tsei³⁵　jong³⁵ tau²²　jet⁴² tsik⁵⁴　kei³⁵ tsei³⁵
this　　CL　　child　　　　　　　feed-CONT　one-CL　　　frog

養到　　　　　蝲仔，　　　咁　　在　　嗰隻　　　□　　去　　放緊
iɔŋ³⁵ tau²²　kei³⁵ tsɐi³⁵　kɐn²²　tsʰui⁵⁴ kɔ³⁵ tsik⁵⁴　aŋ⁴²　hi²²　fɔŋ²² kɐn³⁵
feed-CONT　frog　　　　　then　in　　that-CL　　jar　go　　put-CONT

似頭。　　　小朋友　　　　　　就　　　訓後　　　　　　覺，　嗰隻
tsʰi²² tʰɐu³¹　siu³⁵ pʰɐŋ³¹ jɐu³⁵　tsʰɐu⁵⁴ fɐn²² hɐu⁵⁴　　kau²²　kɔ³⁵ tsik⁵⁴
there　　　　child　　　　　　then　sleep-PFV　　　　sleep　that-CL

青蛙　　　　　就　　走　　出來，　　就　　在　　窗眼
tsʰiaŋ⁴² wa⁴²　tsʰɐu⁵⁴ tsɐu³⁵　tsʰɐt⁴² lui³¹　tsʰɐu⁵⁴ tsʰui⁵⁴ tsʰɔŋ⁴² ŋan²²
frog　　　　　then　walk　out　　　then　at　　window

走後。　　　咁　　嗰　　小朋友　　　　　　隨□　　　　　　去　　搵
tsɐu³⁵ hɐu⁵⁴　kɐn²²　kɔ³⁵　siu³⁵ pʰɐŋ³¹ jɐu³⁵　tsʰui³¹ tʰiaŋ⁵⁴　　hi²²　wɐn³⁵
leave-PFV　then　that　child　　　　　　　everywhere　　　go　　find

嗰條　　　　狗　　　又　　頂　　　　　　翻　　嗰隻　　　□　　去　　搵
kɔ³⁵ tʰiu³¹　kɐu³⁵　jɐu⁵⁴　tɐŋ³⁵　　　fan⁴²　kɔ³⁵ tsik⁵⁴　aŋ⁴²　hi²²　wɐn³⁵
that-CL　　dog　　again　put head into　back　that-CL　　jar　to　　find

又　　搵　　唔　　到。　跟尾　　　就　　嗰　　小朋友
jɐu⁵⁴　wɐn³⁵　m³¹　tau²²　kɐn⁴² mi²²　tsʰɐu⁵⁴ kɔ³⁵　siu³⁵ pʰɐŋ³¹ jɐu³⁵
also　find　NEG　get　　then　　　next　that　child

走　　出　　　窗眼　　　　　　去　　搵，　嗰條　　　　狗
tsɐu³⁵　tsʰɐt⁴²　tsʰɔŋ⁴² ŋan²²　hi²²　wɐn³⁵　kɔ³⁵ tʰiu³¹　kɐu³⁵
walk　out　　window　　　　　to　　find　that-CL　　dog

在　　屋企　　　頂　　　落　隻　□　去，　就　連
tsʰui⁵⁴ uk⁴² kʰi²²　tɐŋ³⁵　lɔk⁵⁴ tsik⁵⁴ an⁴² hi²² tsʰɐu⁵⁴ lin³¹
in　　home　　　put head　into　CL　jar　go　then　together

隻　□　都　□　　　　　　出　嗰　窗眼
tsik⁵⁴ aŋ⁴² tɐu⁴² tuŋ³⁵　　　tsʰɐt⁴² kɔ³⁵ tsʰɔŋ⁴² ŋan²²
CL　jar　both　stretch forward　out　that　window

裏　去　搵，　嗰條　　狗　就　□　落去　　就　打
li²² hi²² wɐn³⁵ kɔ³⁵ tʰiu³¹ kɐu³⁵ tsʰɐu⁵⁴ tsuŋ⁴² lɔk⁵⁴ hi²²　tsʰɐu⁵⁴ ta³⁵
inside go find that-CL dog then smash down　then hid

爛　隻　□。　咁　嗰　小朋友　　　　落　　到
lan⁵⁴ tsik⁵⁴ aŋ⁴² kɐn²² kɔ³⁵ siu³⁵ pʰɐŋ³¹ jɐu³⁵ l　ɔk⁵⁴　tau²²
break CL jar Then that child　　　　descend　to

嗰　窗眼　　　下來，　就　好　焗氣　　就
kɔ³⁵ tsʰɔŋ⁴² ŋan²²　ha⁵⁴ lui³¹ tsʰɐu⁵⁴ hau³⁵ kuk⁴² hi²² tsʰɐu⁵⁴
that window　down　then very angry　then

攔緊　　　　嗰條　　狗　來。　咁　嗰　細蚊仔
lan³¹ kɐn³⁵　kɔ³⁵ tʰiu³¹ kɐu³⁵ lui³¹ kɐn²² kɔ³⁵ sɐi²² mɐn⁴² tsɐi³⁵
embrace-PROG　that-CL dog come Then that child

同　嗰　狗　隨□　　　去　搵，　隨□　　　去　喊。
tʰuŋ³¹ kɔ³⁵ kɐu³⁵ tsʰui³¹ tʰiaŋ⁵⁴ hi²² wɐn³⁵ tsʰui³¹ tiaŋ⁵⁴ hi²² ham²²
and that dog everywhere go find everywhere to shout

咁　在　嗰□　　　就　睇到　一寶　蛾蜂，
kɐn²² tsʰui⁵⁴ kɔ³⁵ tʰiaŋ⁵⁴ tsʰɐu⁵⁴ tʰɐi³⁵ tau²² jɐt⁴² tɐu²² ŋɔ³¹ fuŋ⁴²
then at that place then see one-CL wasp

嗰條　　狗　就　走　去　喊　走　去　吼。
kɔ³⁵ tʰiu³¹ kɐu³⁵ tsʰɐu⁵⁴ tsɐu³⁵ hi²² ham²² tsɐu³⁵ hi²² hɐu⁴²
that-CL dog then go to shout go to bark

227

吼　　之後　　　　咧，　嗰　　啲　　蛾蜂　　　　就　　　　走　　向
hɐu⁴²　tsi⁴² hɐu⁵⁴　lɛ²²　kɔ³⁵　ti⁵⁴　ŋɔ³¹ fuŋ⁴²　tsʰɐu⁵⁴ tsɐu³⁵　hiɔŋ⁴²
bark　after　　　　PRT　that　some　wasp　　　　then　　go　　towards

走　　上　　嗰蔸　　　　樹。　嗰　　細蚊仔　　　　　　跌　　落來，
tsɐu³⁵　sɔŋ⁵⁴　kɔ³⁵ tɐu⁴²　si⁵⁴　kɔ³⁵　sɐi²² mɐn⁴² tsɐi³⁵　tit⁵⁴　lɔk⁵⁴ lui³¹
go　　up　　that-CL　　tree　that　child　　　　　　　fall　down

咁　　跟尾　　　　嗰竇　　　　蛾蜂　　　　對到　　　　嗰條
kɐn²²　kɐn⁴² mi²²　kɔ³⁵ tɐu⁴²　ŋɔ³¹ fuŋ⁴²　tui²² tau²²　kɔ³⁵ tʰiu³¹
then　next　　　　that-CL　　wasp　　　　face-PROG　that-CL

狗仔　　　　跟緊　　　　去。　跟緊去　　　　　咧，
kɐu³⁵ tsɐi³⁵　kɐn⁴² kɐn³⁵　hi²²　kɐn⁴² kɐn³⁵ hi²²　lɛ²²
dog-DIM　　follow-PROG　go　follow-PROG　go　PRT

咁　　跟尾　　　　原先　　　　嗰蔸　　　　樹　　有　　隻　　崖婆，
kɐn²²　kɐn⁴² mi²²　in³¹ sin⁴²　kɔ³⁵ tɐu⁴²　si⁵⁴　jɐu²²　tsik⁵⁴　ŋai³¹ pʰɔ³¹
then　next　　　　original　　that-CL　　tree　have　CL　owl

嗰個　　　　崖婆　　　　在　　嗰蔸　　　　樹　　嗰隻　　　　空
kɔ³⁵ kɔ²²　ŋai³¹ pʰɔ³¹　tsʰui⁵⁴　kɔ³⁵ tɐu⁴²　si⁵⁴　kɔ³⁵ tsik⁵⁴　kɔŋ⁴²
that-CL　　owl　　　　in　　that-CL　　tree　that-CL　　hollow

嗰□　　　　飛　　出來，　　嗰隻　　　　細蚊仔　　　　就
kɔ³⁵ tʰiaŋ⁵⁴　fi⁴²　tsʰɐt⁴² lui³¹　kɔ³⁵ tsik⁵⁴　sɐi²² mɐn⁴² tsɐi³⁵　tsʰɐu⁵⁴
that place　fly　out　　　　that-CL　　child　　　　　then

嚇　　到。　嚇　　到　　咁　　跟尾　　　　嗰　　細蚊仔　　　　　　就
ha²² tau²²　ha²²　tau²²　kɐn²²　kɐn⁴² mi²²　kɔ³⁵　sɐi²² mɐn⁴² tsɐi³⁵　tsʰɐu⁵⁴
scare　get　scare　get　then　next　　　　that　child　　　　　　　then

又　　隨□　　　　　去搵，　　就　　在　　楞高　　　　嗰□
jɐu⁵⁴　tsʰui³¹ tʰiaŋ⁵⁴　hi²² wɐn³⁵　tsʰɐu⁵⁴　tsʰui⁵⁴　ɐŋ⁵⁴ kau⁴²　kɔ³⁵ tʰiaŋ⁵⁴
again　everywhere　go　find　then　at　high-place　that　place

228

原先	嗰蔸	樹	嗰□		就	係	一隻	
in³¹ sin⁴²	kɔ³⁵ tɐu⁴²	si⁵⁴	kɔ³⁵ tʰiaŋ⁵⁴		tsʰɐu⁵⁴	hɐi⁵⁴	jɐt⁴² tsik⁵⁴	
original	that-CL	tree	that	place	then	be	one-CL	

鹿角	來嘅，	□到	一隻		鹿。	咁	嗰隻
luk⁵⁴ kɔk⁴²	ui³¹ kɐ²²	piaŋ⁴² tau²²	jɐt⁴² tsik⁵⁴		luk⁵⁴	kɐn²²	kɔ³⁵ tsik⁵⁴
deer horn	PRT	hide-CONT	one-CL		deer	then	that-CL

鹿	一	升	起	頭	來	唎，	嗰隻	人	嚇	到，
luk⁵⁴	jɐt⁴²	sɐŋ⁴²	hi³⁵	tʰɐu³¹	lui³¹	lɐ²²	kɔ³⁵ tsik⁵⁴	jɐn³¹	ha²²	tau²²
deer	once	rise	up	head	come	PRT	that-CL	man	scare	get

嗰條	狗仔	跟	在	鹿	嗰□	下低
kɔ³⁵ tʰiu³¹	kɐu³⁵ tsɐi³⁵	kɐn⁴²	tsʰui⁵⁴	luk⁵⁴	kɔ³⁵ tʰiaŋ⁵⁴	ha⁵⁴ tɐi⁴²
that-CL	dog-DIM	follow	at	deer	that place	down

嗰□	就	在頭	喊。	嗰條	鹿	唎，
kɔ³⁵ tʰiaŋ⁵⁴	tsʰɐu⁵⁴	tsʰui⁵⁴ tʰɐu³¹	ham²² 。	kɔ³⁵ tʰiu³¹	luk⁵⁴	lɐ²²
that place	then	over there	shout	that-CL	deer	PRT

嗰隻	細蚊仔	騎緊	嗰隻	鹿角，
kɔ³⁵ tsik⁵⁴	sɐi²² mɐn⁴² tsɐi³⁵	kʰiɛ³¹ kɐn³⁵	kɔ³⁵ tsik⁵⁴	luk⁵⁴ kɔk⁴²
that-CL	child	ride-PROG	that-CL	deer horn

嗰條	鹿	咁	□緊	佢	走。	嗰條	狗仔
kɔ³⁵ tʰiu³¹	luk⁵⁴	kɐn²²	piɛ²²kɐn³⁵	kʰi³⁵	tsɐu³⁵	kɔ³⁵ tʰiu³¹	kɐu³⁵ tsɐi³⁵
that-CL	deer	then	carry-PROG	3ʳᵈ SG	go	that-CL	dog-DIM

跟緊	嗰條	鹿	吼，	一路	吼	一路	吼。
kɐn⁴² kɐn³⁵	kɔ³⁵ tʰiu³¹	luk⁵⁴	hɐu⁴²	jɐt⁴² lu⁵⁴	ɐu⁴²	jɐt⁴² lu⁵⁴	hɐu⁴²
follow-PROG	that-CL	deer	bark	keep	bark	keep	bark

咁	嗰條	鹿	唎	去	到	嗰□	嗰間
kɐn²²	kɔ³⁵ tʰiu³¹	luk⁵⁴	lɐ²²	hi²²	tau²²	kɔ³⁵ tʰiaŋ⁵⁴	kɔ³⁵ kan⁴²
then	that-CL	deer	PRT	go	arrive	that place	that-CL

229

屋頂　　　　嗰□　　　　　唎，　　嗰　　細蚊仔　　　　　　俾
uk⁴² tɐŋ³⁵　kɔ³⁵ tʰiaŋ⁵⁴　lɐ²²　kɔ³⁵　sɐi²² mɐn⁴² tsɐi³⁵　pi³⁵
roof　　　　that place　　PRT　　that　child　　　　　　　PASS

嗰條　　　　鹿　　掀　　落去，　　連　　嗰條　　　　狗仔
kɔ³⁵ tʰiu³¹　luk⁵⁴　hin⁴²　lɔk⁵⁴ hi²²　lin³¹　kɔ³⁵ tʰiu³¹　kɐu³⁵ tsɐi³⁵
that-CL　　deer　throw　down　　　even　that-CL　　dog-DIM

都　　跌　　落　　嗰　　地下　　　去。　　咁　　之後　　　　隻
tɐu⁴²　tit⁵⁴　lɔk⁵⁴　kɔ³⁵　tʰi⁵⁴ ha²²　hi²²　kɐn²²　tsi⁴² hɐu⁵⁴　tsik⁵⁴
also　fall　down　that　ground　　go　　then　after　　　CL

細蚊仔　　　　　　俾　　佢　　跌　　落去　　　後　　唎，
sɐi²² mɐn⁴² tsɐi³⁵　pi³⁵　kʰi³⁵　tit⁵⁴　lɔk⁵⁴ hi²²　hɐu⁵⁴　lɐ²²
child　　　　　　PASS　3ʳᵈ SG　fall　down　　after　PRT

咁　　嗰條　　　　狗仔　　　　又　　隨□　　　　　在　　□　　　　搵，
kɐn²²　kɔ³⁵ tʰiu³¹　kɐu³⁵ tsɐi³⁵　jɐu⁵⁴　tsʰui³¹ tʰiaŋ⁵⁴　tsʰui⁵⁴　liaŋ⁵⁴　wɐn³⁵
then　that-CL　　dog-DIM　　again　everywhere　　at　　that place　find

又　　在　　□　　　　樹窿空　　　　去　　搵，　　嗰　　狗仔　　　　又
jɐu⁵⁴　tsʰui⁵⁴　liaŋ⁵⁴　si⁵⁴ luŋ³¹ kɔŋ⁴²　hi²²　wɐn³⁵　kɔ³⁵　kɐu³⁵ tsɐi³⁵　jɐu⁵⁴
also　at　　(that) place　tree　hollow　go　find　　that　dog-DIM　　also

似頭　　　　　搵。　嗰　　狗仔　　　　在　　嗰　　樹窿空
tsʰi²² tʰɐu³¹　wɐn³⁵　kɔ³⁵　kɐu³⁵ tsɐi³⁵　tsʰui⁵⁴　kɔ³⁵　si⁵⁴ luŋ³¹ kɔŋ⁴²
over there　　find　that　dog-DIM　at　　that　tree hollow

嗰　　樹邊　　　　就　　睇到　　　兩隻　　　蜗仔，
kɔ³⁵　si⁵⁴ pin⁴²　tsʰɐu⁵⁴　tʰɐi³⁵ tau²²　liɔŋ³⁵ tsik⁵⁴　kɐi³⁵ tsɐi³⁵
that　tree side　then　see　　　two-CL　　frog

睇到　　　　兩隻　　　我哋　　　喊　　蟾蜍。　　　　　嗰隻
tʰɐi³⁵ tau²²　liɔŋ³⁵ tsik⁵⁴　ŋɔ³⁵ ti⁴²　ham²²　kɐm³¹ si³¹ kɔ³⁵　tsik⁵⁴
see　　　　two-CL　　2ⁿᵈnPL　call　toad　　　　　　hat-CL

狗仔	睇到	一隻		就	好	歡喜	就
kɐu³⁵ tsɐi³⁵	tʰɐi³⁵	tau²² jɐt⁴² tsik⁵⁴		tsʰɐu⁵⁴	hau³⁵	fun⁴² hi³⁵	tsʰɐu⁵⁴
dog-DIM	see	one-CL		then	very	happy	then

似頭	喊	佢，	咁	嗰	細蚊仔		可能	尋
tsʰi²² tʰɐu³¹	ham²²	kʰi³⁵	kɐn²²	kɔ³⁵	sɐi²² mɐn⁴² tsɐi³⁵		hɔ³⁵ nɐŋ³¹	tsʰɐm³¹
over there	call	3ʳᵈ SG	then	CL	child		perhaps	find

去劮，	尋	去	劮	佢	就	好似	蔫後	咁，
hi²² kʰui⁵⁴	tsʰɐm³¹	hi²²	kʰui⁵⁴	kʰi³⁵	tsʰɐu⁵⁴	hau³⁵ tsʰi²²	in²² hɐu⁵⁴	kɐn²²
to tired	find	to	tired	3ʳᵈ SG	then	seemingly	wither-PFV	so

就	撲緊	嗰兜	樹，	似頭	好似	訓覺
tsʰɐu⁵⁴	pʰuk⁴² kɐn³⁵	kɔ³⁵ tɐu⁴²	si⁵⁴	tsʰi²² tʰɐu³¹	hau³⁵ tsʰi²²	fɐn²² kau²² k
then	lean-CONT	that-CL	tree	over there	seemingly	sleep

咁。	咁	嗰	細蚊仔		咧	咁	嗰條	狗仔
ɐn²²	kɐn²²	kɔ³⁵	sɐi²² mɐn⁴² tsɐi³⁵		lɐ²²	kɐn²²	kɔ³⁵ tʰiu³¹	kɐu³⁵ tsɐi³⁵
so	then	that	child		PRT	then	that-CL	dog-DIM

就	似頭	喊，	似頭	喊。	咁	嗰	細蚊仔
tsʰɐu⁵⁴	tsʰi²² tʰɐu³¹	ham²²	tsʰi²² tʰɐu³¹	ham²²	kɐn²²	kɔ³⁵	sɐi²² mɐn⁴² tsɐi³⁵
then	over there	shout	over there	shout	so	that	child

就	俾	佢		吵	醒	後，	咁	跟尾	嗰
tsʰɐu⁵⁴	pi³⁵	kʰi³⁵		tsʰau³¹	siaŋ³⁵	hɐu⁵⁴	kɐn²²	kɐn⁴² mi²²	kɔ³⁵
then	PASS	3ʳᵈ SG		annoy	awake	after	then	next	that

細蚊仔		就	走後。	咁	就	嗰	兩隻
sɐi²² mɐn⁴² tsɐi³⁵		tsʰɐu⁵⁴	tsɐu³⁵ hɐu⁵⁴	kɐn²²	tsʰɐu⁵⁴	kɔ³⁵	liɔŋ³⁵ tsik⁵⁴
child		then	go-PFV	so	then	that	two-CL

蟾蜍	蜗仔		就	走	到	嗰兜	樹面，
kɐm³¹ si³¹	kɐi³⁵ tsɐi³⁵		tsʰɐu⁵⁴	tsɐu³⁵	tau²²	kɔ³⁵ tɐu⁴²	si⁵⁴ min⁵⁴
toad	frog		then	go	arrive	that-CL	tree surface

231

又　　還有　　　幾隻　　　　蟾蜍仔　　　　　又　　走　　到
jɐu⁵⁴　han³¹ jɐu²²　ki³⁵ tsik⁵⁴　kɐm³¹ si³¹ tsɐi³⁵　jɐu⁵⁴　tsɐu³⁵　tau²²
also　　still have　　some-CL　　toad-DIM　　　　also　go　　arrive

嗰苑　　　　樹筒　　　　楞　　嗰□　　　　去　　同　　佢哋　　　　講，
ko³⁵ tɐu⁴²　si⁵⁴ tʰɔŋ³¹　lɐŋ⁵⁴　ko³⁵ tʰiaŋ⁵⁴　hi²²　tʰuŋ³¹　kʰi³⁵ ti⁵⁴　　kɔŋ³⁵
that-CL　　tree trunk　top　　that place　　to　　with　　3ʳᵈ PL　　talk

同　　條　　狗仔　　　　□埋　　　　嗰隻　　　　細蚊仔
tʰuŋ³¹　tʰiu³¹　kɐu³⁵ tsɐi³⁵　lɐu²²mai³¹　ko³⁵ tsik⁵⁴　sɐi²² mɐn⁴² tsɐi³⁵
with　　CL　　dog-DIM　　mix with　　that-CL　　child

講　　再見。
kɔŋ³⁵　tsui²² kin²²
say　　good-bye

Translation to Standard Chinese:

這個小孩子養了一隻青蛙，就在那個罐子裡面把它放著。小孩子睡了覺，那隻青蛙就走了出來，從窗子跑了。小孩子就到處去找，那條狗又把頭頂進罐子里去找，也找不到。然後小孩子走到窗子外面去找，狗在屋裡頂著那個罐子，就連帶著罐子探出那個窗口去找，（然後）狗就掉下去了，打爛了那個罐子。那個小孩子下到窗戶下面來，很生氣，就抱著那條狗。小孩子和狗就到處去找，到處呼喚（青蛙）。在那個地方就看到一窩蜜蜂，狗就過去叫、過去吠。吠完之後，那些蜜蜂就飛上那棵樹。小孩子掉下來，然後那窩蜜蜂就對著那條小狗，跟著它。跟著它呢，原來那棵樹（上）有隻貓頭鷹，那隻貓頭鷹從樹的那個樹洞那邊飛出來，那個小朋友就（被）嚇著了。嚇著了以後，小朋友就又到處去找，在高處原來那棵樹的地方（有）一隻鹿角，藏著一隻鹿。那隻鹿一抬起頭來呢，那个人（被）嚇著了，小狗跟在鹿的下面在那兒叫。那隻鹿呢……那個小孩子騎著鹿角，那隻鹿就馱著他走。小狗跟著鹿吠，一直吠，吠了一路。那隻鹿呢去到屋頂那個地方，小孩子被鹿掀下去了，連

小狗都跌落到地下去了。後來，那個小孩子被它扔到下面去之後呢，小狗又到處在那兒找（青蛙），又到樹洞裡面去找，小狗又在那兒找。那隻小狗在那個樹洞，在那樹邊就看到兩隻青蛙，看到兩隻……我們叫（它）蟾蜍。那隻小狗看到一隻就好高興，就在那兒叫它，那個小朋友可能找累了，找累了它就好像蔫了似的，就靠著那棵樹，在那兒好像睡覺似的。那個小朋友呢，那條小狗就在那兒叫，在那兒叫。小朋友就被它吵醒了，後來小朋友就走了。那兩隻蟾蜍，青蛙就走到樹面上來，另外還有幾隻小蟾蜍又走到樹幹哪兒去跟他們說，跟那條小狗和那個小朋友說再見。

Translation to English:

A child is raising a frog. He puts it into a jar. After the child falls asleep, the frog comes out and leaves through the window. Then the child looks for the frog everywhere. A dog also puts its head into the jar to look for the frog but cannot find it. Then the child goes out of the window and looks for (the frog) outside. The dog has the jar on its head and stretches out of the window to look for (the frog). (Then) it falls down and breaks the jar. The child comes down (to the ground) underneath the window and is hungry, so he holds the dog in his arms. The child and the dog then look for the frog everywhere and call for it everywhere. Over there they see a swarm of wasps, so the dog goes to bark (at the wasps). After that, those wasps fly up to the tree, and the child falls down. Next, the swarm of wasps follow the dog (in its direction). There is an owl in the tree, and the owl flies out of the tree hollow. The kid is scared. Then he again looks for (the frog) everywhere. There is a deer horn in the high place where the tree is, and a deer is hiding there. The deer lifts up its head, and that person (the child) is scared. The dog follows the deer and barks underneath the deer. The deer… The child is riding on the deer horn, and the deer is carrying him around. The dog is following the deer and keeps on barking. The deer goes to the roof place, and the child is thrown off by the deer. Together with him the dog also falls down onto the ground. Later, after the child is thrown off by it (the deer), the dog looks for (the frog) over there (on the ground) and in the tree hollow. The dog again looks for (the frog) over there. In the tree hollow… beside the tree the dog sees two

233

frogs, two (what) we call toads. The dog sees a frog so it is very happy, and it calls him. The child may be tired from searching (for the frog), and he seems exhausted. So he leans on the tree, and he seems asleep over there. So the child… the dog is barking, and then the child is woken up by it. After that the child leaves. The two toads… frogs then go to the tree surface. Some other little toads also go to the trunk of the tree and say good-bye to the dog and the child.

Appendix C: Homophonous Morphemes

As a complement of the syllabary, this section sorts out all Dapeng morphemes according to their phonetic forms.

(1) [-a]

[p-]	42	爸 巴 芭 疤	
	31		
	35	把	
	22	霸 壩	
	54		
[pʰ-]	42		
	31	爬 琶	
	35		
	22	怕	
	54	罷	
[f-]	42	花	
	31		
	35		
	22	化	
	54		
[m-]	42		
	31	麻	
	35	馬	
	22	碼 嫲	
	54	罵	
[w-]	42	蛙 洼	
	31	華 話	
	35		
	22		
	54	畫	
[t-]	42		
	31		
	35	打	
	22		
	54		
[tʰ-]	42	他	
	31		
	35		
	22		

	54		
[n-]	42		
	31	拿	
	35	嫲	
	22		
	54		
[l-]	42		
	31		
	35		
	22		
	54		
[ts-]	42	抓 渣	
	31		
	35	者	
	22		
	54	蔗 炸 詐 榨	
[tsʰ-]	42	車 差	
	31	茶 搽 查	
	35		
	22		
	54	乍	
[s-]	42	沙 紗	
	31	蛇	
	35	灑 耍	
	22	社 捨	
	54	射	
[j-]	42		
	31		
	35		
	22		
	54		

236

[k-]	42	加 瓜 家 嘉
	31	
	35	寡 假
	22	架 嫁 掛 價 稼 駕 卦
	54	
[kʰ-]	42	垮 跨
	31	
	35	
	22	
	54	
[ŋ-]	42	
	31	牙 娃
	35	瓦 雅 訝
	22	
	54	芽
[h-]	42	蝦
	31	霞
	35	
	22	嚇
	54	夏 下 廈 暇 瑕
Zero	42	啞 亞 丫 鴉 阿
	31	
	35	
	22	
	54	

(2) [-ai]

[p-]	42	
	31	
	35	擺
	22	拜
	54	
[pʰ-]	42	
	31	排 牌
	35	
	22	派
	54	敗
[f-]	42	
	31	
	35	
	22	快 塊 筷
	54	
[m-]	42	

	31	埋
	35	
	22	買
	54	賣
[w-]	42	歪
	31	懷
	35	
	22	
	54	壞
[t-]	42	
	31	
	35	
	22	帶 戴
	54	
[tʰ-]	42	
	31	
	35	
	22	太 態
	54	大
[n-]	42	
	31	
	35	乃
	22	奶
	54	耐
[l-]	42	
	31	
	35	
	22	
	54	拉 賴
[ts-]	42	齋 災 栽
	31	
	35	
	22	債
	54	
[tsʰ-]	42	差 釵 猜 搓
	31	柴
	35	踩
	22	蔡
	54	寨

[s-]	42	□
	31	
	35	
	22	逝 誓 曬
	54	
[j-]	42	
	31	椰
	35	□
	22	
	54	
[k-]	42	街 歸 芥 乖 皆 階 偕 佳
	31	
	35	
	22	介 怪 戒 尬 屆 界 疥
	54	
[kʰ-]	42	
	31	
	35	
	22	
	54	
[ŋ-]	42	
	31	崖
	35	
	22	
	54	艾 涯
[h-]	42	
	31	鞋 孩
	35	
	22	
	54	械 □
Zero	42	唉 埃
	31	
	35	矮
	22	
	54	

(3)　[-au]

[p-]	42	包 胞
	31	
	35	飽 保 寶 堡
	22	報

	54	鮑
[pʰ-]	42	
	31	刨 袍
	35	跑
	22	抱 豹 爆 炮 泡
	54	暴
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	毛 茅
	35	
	22	
	54	帽 貌 貓 冒
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	刀
	31	
	35	倒 島
	22	到
	54	
[tʰ-]	42	滔
	31	桃 投 濤 逃 陶 掏 □
	35	討 導
	22	稻 套
	54	盜 道
[n-]	42	
	31	
	35	腦 惱
	22	
	54	鬧 □
[l-]	42	
	31	撈 牢 勞 □
	35	老
	22	
	54	

[ts-]	42	糟 遭
	31	
	35	早 找 棗 爪
	22	罩 灶
	54	
[tsʰ-]	42	抄 鈔 操
	31	曹 巢 售 酬 稠 綢 籌 吵
	35	草 糙 炒
	22	臭 皂 澡 躁 燥
	54	造
[s-]	42	搜 稍 潲 嘯
	31	
	35	
	22	掃 嫂 掃
	54	受 授 壽
[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	高 交 較 郊 膠 餃 羔 膏 糕 蒿
	31	
	35	搞 狡 絞 酵 校 稿
	22	教 窖 覺 教 告
	54	
[kʰ-]	42	
	31	
	35	巧 竅 考 烤 朽
	22	敲 叩 寇 靠
	54	
[ŋ-]	42	
	31	熬
	35	
	22	咬
	54	傲
[h-]	42	
	31	毫 豪
	35	好
	22	孝 好 耗
	54	效 浩 號 號
Zero	42	

		31	
		35	襖
		22	奧 懊
		54	
(4)	[-am]		
	[p-]	42	
		31	
		35	
		22	
		54	
	[pʰ-]	42	
		31	
		35	
		22	
		54	
	[f-]	42	
		31	
		35	
		22	
		54	
	[m-]	42	
		31	
		35	
		22	
		54	
	[w-]	42	
		31	
		35	
		22	
		54	
	[t-]	42	擔
		31	
		35	膽
		22	擔
		54	
	[tʰ-]	42	貪
		31	潭 痰 談 譚
		35	
		22	探
		54	

[n-]	42	
	31	南 男
	35	
	22	
	54	
[l-]	42	
	31	藍 籃
	35	欖
	22	□
	54	覽 攬 纜 濫
[ts-]	42	
	31	
	35	斬 眨
	22	
	54	
[tsʰ-]	42	參
	31	蠶
	35	慘
	22	
	54	站
[s-]	42	三 衫
	31	
	35	
	22	
	54	
[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	尷 甘 柑
	31	
	35	減 敢 感
	22	監 鑒 監 橄
	54	
[kʰ-]	42	
	31	
	35	砍 □
	22	
	54	
[ŋ-]	42	□
	31	岩
	35	

	22	
	54	
[h-]	42	坎
	31	咸 函
	35	
	22	喊
	54	陷
Zero	42	
	31	
	35	
	22	暗
	54	

(5)　[-an]

[p-]	42	班
	31	
	35	板 版 叛
	22	
	54	
[pʰ-]	42	攀
	31	
	35	
	22	
	54	辦 扮 辮 辨 辯 盼
[f-]	42	翻 番
	31	煩 繁 帆 凡
	35	反
	22	販
	54	飯 犯 范 泛 幻
[m-]	42	
	31	蠻
	35	
	22	晚
	54	萬 慢 幔 漫 饅
[w-]	42	彎 灣
	31	還 環 □
	35	挽 玩
	22	
	54	

240

[t-] 42 丹 單
31
35
22 誕 旦
54
[tʰ-] 42 攤 灘
31 彈 壇
35 毯 坦 但
22 淡 炭 嘆
54 蛋 彈
[n-] 42
31 難
35
22
54 難
[l-] 42
31 蘭 攔 欄
35
22
54 爛
[ts-] 42
31
35 盞
22 贊
54
[tsʰ-] 42 餐
31 殘
35 産 鏟
22 燦
54 賺 棧
[s-] 42 山 訕 珊
31
35 散
22 散
54
[j-] 42
31
35
22
54
[k-] 42 間 關 奸 艱
31
35 揀 簡

22 間
54
[kʰ-] 42
31
35
22
54
[ŋ-] 42
31 顏 頑
35
22 眼 研
54 岸 諺 雁
[h-] 42
31 閑
35
22
54 限
Zero 42
31
35
22
54

(6) [-aŋ]
[p-] 42 兵
31
35
22
54
[pʰ-] 42 □
31 棚 彭 膨
35
22 棒
54
[f-] 42
31
35
22
54 患
[m-] 42
31 盲
35
22 孟
54

[w-] 42
31 橫
35
22
54

[t-] 42
31
35
22
54

[tʰ-] 42
31
35
22
54

[n-] 42
31
35
22
54

[l-] 42
31
35 冷
22
54

[ts-] 42 爭 箏 □
31
35
22
54

[tsʰ-] 42 清
31 橙
35
22
54 鄭

[s-] 42 生 聲 笙 甥
31 城 承
35 省
22 腎
54

[j-] 42
31
35
22

54
[k-] 42 耕 更
31
35
22 更
54

[kʰ-] 42
31
35 梗
22
54

[ŋ-] 42
31
35
22
54 硬

[h-] 42 坑 杭 骯
31
35
22
54

Zero 42
31
35
22
54

(7) [-ap]
[p-] 42
31
35
22
54

[pʰ-] 42
31
35
22
54

[f-] 42
31
35
22
54

242

[m-]	42		[j-]	42	
	31			31	
	35			35	
	22			22	
	54			54	
[w-]	42		[k-]	42	夾 甲
	31			31	
	35			35	
	22			22	
	54			54	狹
[t-]	42	搭 答	[kʰ-]	42	
	31			31	
	35			35	
	22			22	
	54			54	
[tʰ-]	42	塔 塌 榻	[ŋ-]	42	
	31			31	
	35			35	
	22			22	
	54			54	
[n-]	42		[h-]	42	
	31			31	
	35			35	
	22			22	
	54			54	合 盒 洽
[l-]	42		Zero	42	鴨 押
	31			31	
	35			35	
	22			22	
	54	臘 蠟		54	
[ts-]	42				

(8)　[-at]

[p-]	42	八
	31	
	35	
	22	
	54	
[pʰ-]	42	
	31	
	35	
	22	
	54	□

[ts-]	42	
	31	
	35	
	22	
	54	
[tsʰ-]	42	插
	31	
	35	
	22	
	54	雜 集 輯 匣
[s-]	42	
	31	
	35	
	22	
	54	什 拾

[f-]	42	發 法			54	閘 疾
	31			[s-]	42	殺
	35				31	
	22				35	
	54	罰			22	
[m-]	42				54	
	31			[j-]	42	
	35				31	
	22				35	
	54	襪			22	
[w-]	42	挖			54	
	31			[k-]	42	刮
	35				31	
	22				35	
	54	滑 猾			22	
[t-]	42				54	
	31			[kʰ-]	42	
	35				31	
	22				35	
	54				22	
[tʰ-]	42				54	
	31			[ŋ-]	42	
	35				31	
	22				35	
	54	達 特			22	
[n-]	42				54	
	31			[h-]	42	
	35				31	
	22				35	
	54				22	
[l-]	42				54	
	31			Zero	42	壓
	35				31	
	22				35	
	54	辣			22	
[ts-]	42	紮 扎 札			54	
	31					
	35			(9)	[-ak]	
	22			[p-]	42	百 伯
	54	蟄			31	
[tsʰ-]	42	刷 察 册 擦			35	
	31				22	
	35				54	陌
	22					

[pʰ-]	42	柏 拍
	31	
	35	
	22	
	54	白 泊
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	
	35	
	22	
	54	麥 抹 □
[w-]	42	
	31	
	35	
	22	
	54	劃 或 惑
[t-]	42	
	31	
	35	
	22	
	54	
[tʰ-]	42	
	31	
	35	
	22	
	54	敵
[n-]	42	
	31	
	35	
	22	
	54	
[l-]	42	
	31	
	35	
	22	
	54	□
[ts-]	42	摘 惻 責
	31	
	35	
	22	
	54	

[tsʰ-]	42	尺 拆 策 赤 測
	31	
	35	
	22	
	54	宅 澤 擇 籍 藉
[s-]	42	
	31	
	35	
	22	
	54	石
[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	格 革 隔
	31	
	35	
	22	
	54	
[kʰ-]	42	
	31	
	35	
	22	
	54	
[ŋ-]	42	
	31	
	35	
	22	
	54	額 逆
[h-]	42	客
	31	
	35	
	22	
	54	
Zero	42	
	31	
	35	
	22	
	54	

(10) [-ia]

				54	
[p-]	42		[ts-]	42	
	31			31	
	35			35	
	22			22	
	54			54	
[pʰ-]	42		[tsʰ-]	42	
	31			31	
	35			35	
	22			22	
	54			54	
[f-]	42		[s-]	42	些
	31			31	
	35			35	寫
	22			22	泄 卸 瀉 薩
	54			54	
[m-]	42		[j-]	42	
	31			31	
	35			35	
	22			22	
	54	□		54	
[w-]	42		[k-]	42	
	31			31	
	35			35	
	22			22	
	54			54	
[t-]	42		[kʰ-]	42	
	31			31	
	35			35	
	22	借		22	
	54			54	
[tʰ-]	42		[ŋ-]	42	
	31	邪 斜		31	
	35			35	
	22			22	
	54	謝		54	
[n-]	42		[h-]	42	
	31			31	
	35			35	
	22			22	
	54			54	
[l-]	42				
	31				
	35				
	22				

246

Zero	42	
	31	爺
	35	野 惹
	22	也
	54	夜 腋

(11) [-iaŋ]

[p-]	42		[l-]	35	
	31			22	
	35	餅		54	
	22	柄 □		42	
	54			31	
[pʰ-]	42			35	
	31	瓶		22	領 嶺
	35			54	□
	22		[ts-]	42	
	54	病		31	
[f-]	42			35	井
	31			22	
	35			54	
	22		[tsʰ-]	42	青
	54			31	晴
[m-]	42			35	請
	31	名 銘		22	
	35			54	净
	22		[s-]	42	腥
	54	命		31	
[w-]	42			35	醒
	31			22	牲 姓
	35			54	
	22		[j-]	42	
	54			31	
[t-]	42	釘 釘		35	
	31			22	
	35			54	
	22		[k-]	42	驚
	54	□		31	
[tʰ-]	42	廳 聽		35	頸
	31			22	鏡
	35	艇		54	
	22		[kʰ-]	42	
	54			31	
[n-]	42			35	
	31			22	
				54	儉
			[ŋ-]	42	
				31	
				35	
				22	
				54	

[h-] 42 輕
31
35
22
54
Zero 42
31 榮 贏
35 影
22
54

(12) [-iak]
[p-] 42 壁
31
35
22
54
[pʰ-] 42 劈
31
35
22
54
[f-] 42
31
35
22
54
[m-] 42
31
35
22
54
[w-] 42
31
35
22
54
[t-] 42
31
35
22
54
[tʰ-] 42 踢 □
31
35

22
54 笛
[n-] 42
31
35
22
54
[l-] 42
31
35
22
54
[ts-] 42
31
35
22
54
[tsʰ-] 42
31
35
22
54
[s-] 42 錫
31
35
22
54
[j-] 42
31
35
22
54
[k-] 42
31
35
22
54
[kʰ-] 42
31
35
22
54 劇

[ŋ-]	42	
	31	
	35	
	22	
	54	
[h-]	42	
	31	
	35	
	22	
	54	
Zero	42	憶
	31	
	35	
	22	
	54	

(13)　[-ɐi]

[p-]	42	
	31	
	35	
	22	
	54	
[pʰ-]	42	批 □
	31	
	35	
	22	
	54	幣 斃
[f-]	42	揮 輝 徽
	31	
	35	
	22	費 肺 弗 沸 廢
	54	
[m-]	42	
	31	
	35	米
	22	
	54	
[w-]	42	慰 威
	31	圍 為 惟 唯 維 遺 違
	35	委 偉 偽 魏
	22	諱 喂 熨 畏 葦 緯
	54	惠 慧 彗 位 胃 猬 謂 衛
[t-]	42	低

	31	
	35	底 抵
	22	帝
	54	
[tʰ-]	42	梯 涕
	31	題 提 啼 蹄 堤
	35	體
	22	替 剃
	54	弟 第 遞 隸
[n-]	42	
	31	泥
	35	
	22	
	54	
[l-]	42	
	31	犁 黎
	35	禮
	22	
	54	麗 例 厲 勵
[ts-]	42	
	31	
	35	
	22	際 濟 擠 祭 製
	54	
[tsʰ-]	42	妻 凄
	31	
	35	齊
	22	砌
	54	
[s-]	42	西 婿
	31	
	35	使 駛 洗
	22	世 勢 細
	54	
[j-]	42	
	31	
	35	
	22	
	54	

249

[k-]	42	雞 圭 閨 稽 魁
	31	
	35	鬼 傀
	22	貴 桂 季 計 繼
	54	
[kʰ-]	42	溪 虧 規 窺
	31	葵 逵 奎 携
	35	啟 毀
	22	契
	54	跪 櫃
[ŋ-]	42	
	31	危
	35	
	22	
	54	毅 巍 藝 囈
[h-]	42	
	31	桅
	35	
	22	
	54	繫 系 奚
Zero	42	
	31	
	35	
	22	
	54	

(14) [-ɐu]

[p-]	42	□
	31	
	35	
	22	
	54	
[pʰ-]	42	
	31	
	35	剖
	22	
	54	
[f-]	42	
	31	浮
	35	否
	22	
	54	埠 阜
[m-]	42	□

	31	謀
	35	某
	22	
	54	謬 拇 茂 貿
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	
	31	
	35	斗 抖 陡 兜 糾
	22	斗 竇
	54	
[tʰ-]	42	偷
	31	頭
	35	
	22	
	54	豆
[n-]	42	
	31	
	35	扭 紐
	22	
	54	
[l-]	42	
	31	流 樓 溜 留 硫 榴 劉
	35	
	22	柳
	54	漏
[ts-]	42	州 周 洲 鄒
	31	
	35	走 酒
	22	皺 咒
	54	
[tsʰ-]	42	抽 秋
	31	囚
	35	丑 醜
	22	湊 獸 嗅
	54	就 袖

[s-]	42	修 收 羞
	31	愁 仇
	35	手 守
	22	秀 瘦 繡
	54	
[j-]	42	幽
	31	油 由 柔 揉 悠 游 猶
		憂 優 尤 郵
	35	友 柚
	22	有 幼 肴 誘
	54	又 右
[k-]	42	鈎 鳩 勾 溝
	31	
	35	九 苟 狗 久 韭
	22	救 究 够 構 購
	54	
[kʰ-]	42	
	31	求 球
	35	
	22	舅
	54	舊
[ŋ-]	42	
	31	牛
	35	偶
	22	藕
	54	
[h-]	42	丘 休 吼
	31	侯 喉 猴
	35	口
	22	厚
	54	後 垢 后 候
Zero	42	歐 謳
	31	
	35	
	22	嘔 漚
	54	

(15) [-ɐm]

[p-]	42	
	31	
	35	
	22	

	54	
[pʰ-]	42	
	31	
	35	
	22	
	54	
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	
	35	
	22	
	54	
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	
	31	
	35	□
	22	
	54	
[tʰ-]	42	
	31	
	35	
	22	
	54	
[n-]	42	
	31	
	35	
	22	
	54	
[l-]	42	
	31	林 淋 臨
	35	
	22	
	54	

251

[ts-]	42	針	
	31		
	35	枕	
	22	浸 寢	
	54		
[tsʰ-]	42	琛 侵	
	31	尋 沉	
	35		
	22		
	54		
[s-]	42	心 參 深 森	
	31		
	35	慎 審 嬸	
	22	滲	
	54	甚	
[j-]	42	音 蔭 陰	
	31	淫	
	35	飲	
	22		
	54	任 壬	
[k-]	42	今	
	31		
	35	錦	
	22	禁	
	54		
[kʰ-]	42	襟	
	31	琴 禽	
	35		
	22		
	54		
[ŋ-]	42		
	31		
	35		
	22		
	54		
[h-]	42		
	31	含 銜	
	35		
	22		
	54		
Zero	42		
	31		
	35		

	22		
	54		
(16)	[-ɐn]		
[p-]	42	奔 殯 彬 賓	
	31		
	35		
	22		
	54		
[pʰ-]	42		
	31		
	35	品	
	22	□	
	54	笨	
[f-]	42	昏 婚 勳 分 紛 勛 熏 薰	
	31	焚 墳 渾 暈	
	35	粉	
	22	奮 糞 訓	
	54	份	
[m-]	42	蚊	
	31	文 聞 民 眠 紋	
	35		
	22		
	54	問	
[w-]	42	溫 瘟	
	31	魂 雲	
	35	穩	
	22		
	54	運	
[t-]	42	墩 敦	
	31		
	35		
	22		
	54		
[tʰ-]	42	吞	
	31	□	
	35		
	22		
	54		

252

[n-]	42	
	31	
	35	□
	22	
	54	
[l-]	42	
	31	鄰 倫 輪
	35	□
	22	
	54	論
[ts-]	42	真 偵 津 竣 珍 貞
	31	
	35	浚 准 準 疹 診 振 鎮
	22	俊 晉 進 震
	54	
[tsʰ-]	42	春 親 □
	31	陳 塵 秦
	35	
	22	蠢
	54	陣
[s-]	42	新 身 申 伸 詢 辛 薪
	31	神 臣 晨 蓴 醇 辰 誠 唇 巡 旬
	35	筍
	22	信 迅 訊
	54	順
[j-]	42	因 姻 殷
	31	人 刃
	35	隱 忍
	22	引 印
	54	閏 潤 韻
[k-]	42	根 軍 君 金 均 鈞 巾 斤 筋 跟
	31	
	35	緊 謹
	22	
	54	郡
[kʰ-]	42	昆
	31	群 芹
	35	滾
	22	近 菌 捆 困
	54	近 裙

[ŋ-]	42	恩
	31	銀
	35	□
	22	
	54	
[h-]	42	欣
	31	痕
	35	很
	22	
	54	恨
Zero	42	
	31	
	35	
	22	
	54	

(17) [-ɐŋ]

[p-]	42	冰 檳 崩 姘
	31	
	35	丙
	22	並
	54	
[pʰ-]	42	拼
	31	平 朋 貧 頻 屏 評
	35	
	22	
	54	聘
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	萌 盟 明 鳴
	35	
	22	
	54	
[w-]	42	
	31	
	35	
	22	尹 允 永 咏 泳
	54	橫

[t-]	42	燈 丁 登 汀
	31	
	35	等 頂 鼎
	22	凳 訂
	54	
[tʰ-]	42	
	31	廷 亭 庭 停 藤 騰
	35	挺
	22	
	54	定 鄧
[n-]	42	
	31	能 寧 寧
	35	
	22	
	54	
[l-]	42	
	31	靈 零 齡 凌 陵 菱 鱗 憐 鈴
	35	
	22	
	54	另 令 伶 楞
[ts-]	42	精 增 蒸 僧 曾 睛 懲 贈 征 正 睜
	31	
	35	整
	22	正 症 證 政
	54	
[tsʰ-]	42	稱
	31	情 呈 程 層 曾
	35	
	22	秤
	54	
[s-]	42	星 升
	31	成 乘
	35	
	22	性 勝 聖
	54	盛 剩
[j-]	42	應 鷹 嬰 櫻 鶯 鸚 英
	31	仁 扔 仍 螢 迎 寅 盈 營
	35	
	22	
	54	認 孕

[k-]	42	經 庚 京 荊
	31	
	35	景 警 境 竟 競 頃
	22	勁 敬 徑
	54	
[kʰ-]	42	鯨
	31	瓊 勤
	35	肯 墾 傾
	22	
	54	
[ŋ-]	42	
	31	
	35	
	22	
	54	
[h-]	42	兄 卿 興
	31	玄 行 形 弦 型 宏 莖 衡
	35	
	22	慶 興
	54	幸 杏
Zero	42	
	31	
	35	
	22	
	54	

(18) [-ɐp]

[p-]	42	
	31	
	35	
	22	
	54	
[pʰ-]	42	
	31	
	35	
	22	
	54	
[f-]	42	
	31	
	35	
	22	
	54	

254

[m-]	42	
	31	
	35	
	22	
	54	
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	
	31	
	35	
	22	
	54	
[tʰ-]	42	
	31	
	35	
	22	
	54	
[n-]	42	
	31	
	35	
	22	
	54	
[l-]	42	泣
	31	
	35	
	22	
	54	立 粒
[ts-]	42	執 汁
	31	
	35	
	22	
	54	
[tsʰ-]	42	緝 揖
	31	
	35	
	22	
	54	習
[s-]	42	濕
	31	
	35	
	22	
	54	十

[j-]	42	
	31	
	35	
	22	
	54	入
[k-]	42	急 鴿 蛤
	31	
	35	
	22	
	54	
[kʰ-]	42	給 汲 級 磕 吸
	31	
	35	
	22	
	54	極 及
[ŋ-]	42	
	31	
	35	
	22	
	54	
[h-]	42	恰
	31	
	35	
	22	
	54	
Zero	42	
	31	
	35	
	22	
	54	

(19) [-ɐt]

[p-]	42	筆 不 畢
	31	
	35	
	22	
	54	
[pʰ-]	42	匹 珀
	31	
	35	
	22	
	54	

255

[f-]	42	佛
	31	
	35	
	22	
	54	伐
[m-]	42	
	31	
	35	
	22	
	54	蜜 物 密 勿
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	
	31	
	35	
	22	
	54	
[tʰ-]	42	
	31	
	35	
	22	
	54	
[n-]	42	
	31	
	35	
	22	
	54	
[l-]	42	
	31	
	35	
	22	
	54	律
[ts-]	42	質
	31	
	35	
	22	
	54	
[tsʰ-]	42	七 出 漆 膝 蝨
	31	
	35	
	22	
	54	侄 轄

[s-]	42	卒 虱 摔 蟀 失 室
	31	
	35	
	22	
	54	實 术
[j-]	42	一 日
	31	
	35	
	22	
	54	
[k-]	42	骨 吉 擊
	31	
	35	
	22	
	54	□
[kʰ-]	42	屈 咳 窟 克
	31	
	35	
	22	
	54	倔
[ŋ-]	42	乞
	31	
	35	
	22	
	54	
[h-]	42	
	31	
	35	
	22	
	54	核
Zero	42	
	31	
	35	
	22	
	54	

(20) [-ɐk]

[p-]	42	北 逼 璧 迫
	31	
	35	
	22	
	54	

[pʰ-]	42	霹 僻
	31	
	35	
	22	
	54	
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	
	35	
	22	
	54	墨 默 □
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	滴 黑 德
	31	
	35	
	22	
	54	
[tʰ-]	42	
	31	
	35	
	22	
	54	
[n-]	42	
	31	
	35	
	22	
	54	
[l-]	42	
	31	
	35	
	22	
	54	力 曆 歷 勒
[ts-]	42	織 積 側 則 即 鯽 績 迹 脊 職
	31	
	35	
	22	

	54	
[tsʰ-]	42	
	31	
	35	
	22	
	54	直 賊 席 植 殖 值
[s-]	42	識 色 昔 式 飾 適 釋 息 熄 媳 析 惜 塞
	31	
	35	
	22	
	54	碩 夕
[j-]	42	益
	31	
	35	
	22	
	54	翼 液 亦 役 疫 譯
[k-]	42	
	31	
	35	
	22	
	54	
[kʰ-]	42	刻
	31	
	35	
	22	
	54	
[ŋ-]	42	厄 扼 軛
	31	
	35	
	22	
	54	
[h-]	42	黑
	31	
	35	
	22	
	54	
Zero	42	
	31	
	35	
	22	
	54	

(21) [-iɛ]

[p-]	42	
	31	
	35	
	22	
	54	
[pʰ-]	42	
	31	
	35	
	22	
	54	
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	
	35	
	22	
	54	
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	
	31	
	35	
	22	
	54	
[tʰ-]	42	
	31	
	35	
	22	
	54	
[n-]	42	
	31	
	35	
	22	
	54	
[l-]	42	
	31	
	35	
	22	
	54	
[ts-]	42	
	31	
	35	姐
	22	這
	54	
[tsʰ-]	42	
	31	
	35	且
	22	
	54	
[s-]	42	
	31	
	35	
	22	
	54	
[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	
	31	
	35	
	22	
	54	
[kʰ-]	42	
	31	□
	35	
	22	
	54	
[ŋ-]	42	
	31	
	35	
	22	
	54	
[h-]	42	
	31	
	35	
	22	
	54	
Zero	42	
	31	
	35	
	22	
	54	

258

(22) [-ɔ]

[p-]	42	波
	31	
	35	
	22	播
	54	
[pʰ-]	42	
	31	婆
	35	
	22	
	54	
[f-]	42	
	31	
	35	火
	22	貨
	54	
[m-]	42	摩 魔
	31	磨
	35	摸
	22	
	54	磨
[w-]	42	窩
	31	和 禾
	35	
	22	
	54	禍
[t-]	42	多
	31	
	35	躲
	22	
	54	
[tʰ-]	42	拖
	31	舵 陀 馱 駝
	35	妥
	22	
	54	惰 墮
[n-]	42	
	31	
	35	
	22	
	54	糯
[l-]	42	

	31	羅 籮 鑼
	35	
	22	
	54	
[ts-]	42	
	31	
	35	左 阻
	22	
	54	
[tsʰ-]	42	初
	31	鋤
	35	楚
	22	坐 錯 銼
	54	助 座
[s-]	42	疏 梳 蔬
	31	傻
	35	所 瑣 鎖
	22	
	54	
[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	歌 哥
	31	
	35	果
	22	過 個
	54	
[kʰ-]	42	棵
	31	
	35	
	22	課
	54	
[ŋ-]	42	
	31	鵝 俄 蛾
	35	我
	22	
	54	餓 臥

259

[h-]	42	
	31	何 荷 河 荷
	35	可 苛
	22	
	54	賀
Zero	42	
	31	
	35	
	22	
	54	

(23) [-ɔŋ]

[p-]	42	幫 邦
	31	
	35	榜 綁
	22	
	54	
[pʰ-]	42	
	31	憑 傍 旁
	35	
	22	
	54	
[f-]	42	方 肪 彷 芳 妨 謊 荒 慌
	31	房 防 蓬 篷
	35	紡 訪 恍 夥
	22	放 仿
	54	
[m-]	42	
	31	忘 忙 芒 茫 蒙 亡
	35	網 妄
	22	
	54	望 夢
[w-]	42	央 殃 汪
	31	黃 王 皇
	35	枉 往
	22	
	54	旺
[t-]	42	當
	31	
	35	黨 擋 董
	22	棟
	54	

[tʰ-]	42	燙 湯
	31	唐 糖 堂 塘 童 瞳 筒 桐
	35	蕩 躺 趟
	22	
	54	
[n-]	42	
	31	曩 農
	35	
	22	
	54	
[l-]	42	
	31	狼 廊 攏 隆 籠
	35	
	22	
	54	朗 浪
[ts-]	42	張 裝 章 髒 樟 椿 莊
	31	
	35	長 總 掌
	22	帳 壯 悵 葬 宗 障 漲 脹 賬
	54	
[tsʰ-]	42	倉 昌 瘡 窗 蒼 囪 匆 蔥 聰
	31	牀 長 腸 場 床 藏 叢
	35	廠 暢 創 搶
	22	唱 倡
	54	撞 丈 仗 杖 狀 藏 □
[s-]	42	傷 桑 霜 孀 商 嗓
	31	常 嘗 償 裳
	35	爽 賞 喪
	22	宋
	54	上 尚
[j-]	42	
	31	
	35	
	22	
	54	

260

[k-]	42	光 綱 岡 缸 剛 崗 肛 江 扛 烘
	31	
	35	講 廣 港
	22	鋼 虹 降 況
	54	汞
[kʰ-]	42	慷 康 腔 哄
	31	逛 狂
	35	孔
	22	抗 礦 炕 曠 擴 控
	54	
[ŋ-]	42	
	31	
	35	
	22	仰 昂
	54	
[h-]	42	糠 空 空
	31	蝗 行 降 洪 鴻
	35	
	22	航
	54	項 巷
Zero	42	
	31	
	35	
	22	
	54	

(24) [-ɔk]

[p-]	42	剝 駁 博 脖
	31	
	35	
	22	
	54	
[pʰ-]	42	訃
	31	
	35	
	22	
	54	薄
[f-]	42	
	31	
	35	
	22	
	54	

[m-]	42	膜 幕
	31	
	35	
	22	
	54	莫 寞
[w-]	42	鍋 握
	31	
	35	
	22	
	54	獲
[t-]	42	
	31	
	35	
	22	
	54	
[tʰ-]	42	托
	31	
	35	
	22	
	54	
[n-]	42	
	31	
	35	
	22	
	54	
[l-]	42	胳
	31	
	35	
	22	
	54	落 烙 樂 洛 絡 駱
[ts-]	42	桌 作 捉 著
	31	
	35	
	22	
	54	
[tsʰ-]	42	着
	31	
	35	
	22	
	54	着 鑿 續 □

[s-]	42	塑 索
	31	
	35	
	22	
	54	
[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	各 郭 國 角 覺 閣
	31	
	35	
	22	
	54	
[kʰ-]	42	確 霍
	31	
	35	
	22	
	54	
[ŋ-]	42	
	31	
	35	
	22	
	54	樂 岳 嶽 鱷
[h-]	42	殼
	31	
	35	
	22	
	54	學 鶴
Zero	42	惡
	31	
	35	
	22	
	54	

(25) [-iɔ]

[p-]	42	
	31	
	35	
	22	
	54	
[pʰ-]	42	
	31	
	35	

	22	
	54	
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	
	35	
	22	
	54	
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	
	31	
	35	朵
	22	
	54	
[tʰ-]	42	□
	31	
	35	
	22	
	54	
[n-]	42	
	31	
	35	
	22	
	54	
[l-]	42	
	31	螺
	35	
	22	
	54	掠 略
[ts-]	42	
	31	
	35	
	22	
	54	

[tsʰ-] 42 嚼
31
35
22
54
[s-] 42
31
35
22
54
[j-] 42
31
35
22
54
[k-] 42
31
35 裹
22
54
[kʰ-] 42
31 茄
35
22
54
[ŋ-] 42
31
35
22
54
[h-] 42 靴
31
35
22
54
Zero 42
31
35
22
54

(26) [-iɔŋ]
[p-] 42
31
35

22
54
[pʰ-] 42
31
35
22
54
[f-] 42
31
35
22
54
[m-] 42
31
35
22
54
[w-] 42
31
35
22
54
[t-] 42
31
35
22
54
[tʰ-] 42
31
35
22
54
[n-] 42
31 娘
35
22
54
[l-] 42
31 良 粱 輌 凉 梁 糧 量
35 兩
22
54 亮 量 諒

263

[ts-]	42	
	31	
	35	蔣 槳 獎
	22	將 醬 漿 將
	54	象 像
[tsʰ-]	42	槍
	31	牆 祥 詳
	35	
	22	
	54	匠
[s-]	42	相 廂 箱
	31	
	35	想
	22	相
	54	
[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	薑 姜 僵 羌
	31	
	35	
	22	
	54	
[kʰ-]	42	疆 匡 筐
	31	強
	35	強
	22	
	54	
[ŋ-]	42	
	31	
	35	
	22	
	54	
[h-]	42	香 向 鄉
	31	
	35	響 享
	22	餉
	54	
Zero	42	秧
	31	揚 羊 洋 陽 楊
	35	養

	22	癢
	54	讓 樣 釀 嚷 壤

(27) [-iɔk]

[p-]	42	
	31	
	35	
	22	
	54	
[pʰ-]	42	
	31	
	35	
	22	
	54	
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	
	35	
	22	
	54	
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	琢
	31	
	35	
	22	
	54	
[tʰ-]	42	
	31	
	35	
	22	
	54	
[n-]	42	
	31	
	35	
	22	
	54	

[l-]	42			Zero	42	約 躍
	31				31	
	35				35	
	22				22	
	54				54	藥 若 諾 弱
[ts-]	42	爵 雀 卓 啄		(28) [-i]		
	31			[p-]	42	悲 碑 跛 坡
	35				31	
	22				35	比 臂
	54				22	閉 痹 泌 秘
[tsʰ-]	42	戳			54	
	31			[pʰ-]	42	
	35				31	皮 疲
	22				35	鄙 庇 彼 卑 婢 陛 稗 披
	54				22	被
[s-]	42	削			54	備 鼻
	31			[f-]	42	飛 灰 非 菲 妃
	35				31	肥
	22				35	匪
	54				22	
[j-]	42				54	
	31			[m-]	42	
	35				31	微 眉 媚 楣 迷 薇
	22				35	
	54				22	尾 美
[k-]	42	腳			54	未 味 謎 昧
	31			[w-]	42	
	35				31	
	22				35	
	54				22	
[kʰ-]	42	却			54	
	31			[t-]	42	□
	35				31	
	22				35	
	54				22	
[ŋ-]	42				54	
	31			[tʰ-]	42	
	35				31	
	22				35	
	54				22	
[h-]	42				54	地
	31					
	35					
	22					
	54					

265

[n-] 42
　　31 宜 誼
　　35 女 呢
　　22 你 爾
　　54 昵 尼
[l-] 42
　　31 梨 離 璃 籬
　　35 李
　　22 呂 里 理 鯉 厘 侶 旅
　　54 利 荔 慮 濾
[ts-] 42 豬 資 支 知 之 諸 姿 枝 肢 芝 朱 珠 株 蛛 誅 滋
　　31
　　35 紙 主 子 紫 指 脂 止 址 趾 煮
　　22 注 智 至 致 志 痣 鑄 置 駐
　　54
[tsʰ-] 42 咨 雌
　　31 徐 除 遲 祠 飼 馳 詞 匙 池 持 厨 儲 瓷 糍 慈 磁 臍 玆 嗣 辭
　　35 此 齒 處 恥 取 娶 始
　　22 似 柱 處 刺 次 廁 措 翅
　　54 寺 住 字 自 恃 峙 痔 治
[s-] 42 詩 書 絲 師 司 獅 尸 施 舒 私 斯 厮 思 需
　　31 時 署 薯
　　35 死 鼠 史 屎 暑
　　22 試 四 市 輸 肆 賜
　　54 是 士 事 侍 示 樹 視 氏 豉
[j-] 42
　　31
　　35
　　22
　　54
[k-] 42 機 几 肌 基 拘 居 饑
　　31

　　35 舉 幾 紀 杞 豈
　　22 鋸 寄 記 己 句 據 既
　　54
[kʰ-] 42 區 俱 矩 駒 驅
　　31 旗 棋 岐 奇 騎 期 其
　　35 姬
　　22 企
　　54 懼 具 技 忌 巨 拒 距 渠
[ŋ-] 42
　　31
　　35 耳 餌
　　22
　　54 二 儀
[h-] 42 虛 墟 犧 熙 嘻 嬉 噓 希 稀
　　31
　　35 喜 起 許 潊
　　22 去 戲 氣 器 汽 棄 怯
　　54
Zero 42 衣 伊 醫 依
　　31 移 而 魚 餘 兒 如 疑 娛 愚 漁 夷 姨 怡 愈 喻 愉 榆 余
　　35 椅 語 議 己 與 宇 羽 禹
　　22 以 意 雨 于 於
　　54 異 義 蟻 寓 遇 易 預 豫

(29) [-im]
[p-] 42
　　31
　　35
　　22
　　54
[pʰ-] 42
　　31
　　35
　　22
　　54

266

[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	
	35	
	22	
	54	
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	
	31	
	35	點
	22	店 掂
	54	□
[tʰ-]	42	舔 添
	31	甜
	35	
	22	
	54	
[n-]	42	粘
	31	嚴 閹
	35	
	22	染
	54	撚 念 驗
[l-]	42	
	31	廉 鐮
	35	
	22	臉 斂
	54	□
[ts-]	42	尖 瞻 佔
	31	
	35	
	22	占 沾
	54	
[tsʰ-]	42	簽 籤
	31	
	35	
	22	

	54	暫 慚
[s-]	42	
	31	禪 蟬
	35	閃
	22	
	54	
[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	
	31	嫌
	35	檢
	22	劍
	54	
[kʰ-]	42	欽
	31	兼 鉗
	35	
	22	
	54	
[ŋ-]	42	
	31	
	35	
	22	
	54	
[h-]	42	歉 謙
	31	
	35	險
	22	欠
	54	
Zero	42	淹
	31	鹽 炎
	35	掩
	22	厭
	54	焰 艷

(30) [-in]

[p-]	42	邊 編 鞭
	31	
	35	
	22	變
	54	

<table>
<tr><td>[pʰ-]</td><td>42</td><td>偏 遍 篇</td></tr>
<tr><td></td><td>31</td><td></td></tr>
<tr><td></td><td>35</td><td></td></tr>
<tr><td></td><td>22</td><td>片 騙</td></tr>
<tr><td></td><td>54</td><td>便</td></tr>
<tr><td>[f-]</td><td>42</td><td></td></tr>
<tr><td></td><td>31</td><td></td></tr>
<tr><td></td><td>35</td><td></td></tr>
<tr><td></td><td>22</td><td></td></tr>
<tr><td></td><td>54</td><td></td></tr>
<tr><td>[m-]</td><td>42</td><td></td></tr>
<tr><td></td><td>31</td><td>棉 綿</td></tr>
<tr><td></td><td>35</td><td>勉 敏 憫 皿</td></tr>
<tr><td></td><td>22</td><td>緬</td></tr>
<tr><td></td><td>54</td><td>面 麵</td></tr>
<tr><td>[w-]</td><td>42</td><td></td></tr>
<tr><td></td><td>31</td><td></td></tr>
<tr><td></td><td>35</td><td></td></tr>
<tr><td></td><td>22</td><td></td></tr>
<tr><td></td><td>54</td><td></td></tr>
<tr><td>[t-]</td><td>42</td><td>端 奠</td></tr>
<tr><td></td><td>31</td><td></td></tr>
<tr><td></td><td>35</td><td>短</td></tr>
<tr><td></td><td>22</td><td>墊</td></tr>
<tr><td></td><td>54</td><td></td></tr>
<tr><td>[tʰ-]</td><td>42</td><td>天</td></tr>
<tr><td></td><td>31</td><td>田 填 團</td></tr>
<tr><td></td><td>35</td><td></td></tr>
<tr><td></td><td>22</td><td>斷</td></tr>
<tr><td></td><td>54</td><td>殿 電 段 緞 鍛</td></tr>
<tr><td>[n-]</td><td>42</td><td>□</td></tr>
<tr><td></td><td>31</td><td>年</td></tr>
<tr><td></td><td>35</td><td></td></tr>
<tr><td></td><td>22</td><td>暖</td></tr>
<tr><td></td><td>54</td><td>嫩</td></tr>
<tr><td>[l-]</td><td>42</td><td></td></tr>
<tr><td></td><td>31</td><td>連 蓮 聯</td></tr>
<tr><td></td><td>35</td><td>卵</td></tr>
<tr><td></td><td>22</td><td></td></tr>
<tr><td></td><td>54</td><td>亂 煉 練 戀</td></tr>
<tr><td>[ts-]</td><td>42</td><td>專 磚 遵 煎 尊</td></tr>
<tr><td></td><td>31</td><td></td></tr>
<tr><td></td><td>35</td><td>展 揣 剪 轉</td></tr>
<tr><td></td><td>22</td><td>箭 鑽 鑽 戰</td></tr>
<tr><td></td><td>54</td><td></td></tr>
<tr><td>[tsʰ-]</td><td>42</td><td>遷 千 川 穿 殲 村 纖</td></tr>
<tr><td></td><td>31</td><td>全 傳 前 錢 存 痊</td></tr>
<tr><td></td><td>35</td><td>喘 踐 淺</td></tr>
<tr><td></td><td>22</td><td>寸 串 襯 懺 囤 濺</td></tr>
<tr><td></td><td>54</td><td>賤 傳 旋</td></tr>
<tr><td>[s-]</td><td>42</td><td>酸 孫 先 喧 仙 鮮 宣 □</td></tr>
<tr><td></td><td>31</td><td>船 潛 繩 循</td></tr>
<tr><td></td><td>35</td><td>選 癬 損</td></tr>
<tr><td></td><td>22</td><td>扇 線 楦 遜 蒜 算</td></tr>
<tr><td></td><td>54</td><td>善</td></tr>
<tr><td>[j-]</td><td>42</td><td></td></tr>
<tr><td></td><td>31</td><td></td></tr>
<tr><td></td><td>35</td><td></td></tr>
<tr><td></td><td>22</td><td></td></tr>
<tr><td></td><td>54</td><td></td></tr>
<tr><td>[k-]</td><td>42</td><td>肩 堅</td></tr>
<tr><td></td><td>31</td><td></td></tr>
<tr><td></td><td>35</td><td>捲 卷</td></tr>
<tr><td></td><td>22</td><td>見 建 券</td></tr>
<tr><td></td><td>54</td><td></td></tr>
<tr><td>[kʰ-]</td><td>42</td><td></td></tr>
<tr><td></td><td>31</td><td>權 拳</td></tr>
<tr><td></td><td>35</td><td></td></tr>
<tr><td></td><td>22</td><td></td></tr>
<tr><td></td><td>54</td><td>件 健 倦 鍵 腱</td></tr>
<tr><td>[ŋ-]</td><td>42</td><td></td></tr>
<tr><td></td><td>31</td><td></td></tr>
<tr><td></td><td>35</td><td></td></tr>
<tr><td></td><td>22</td><td></td></tr>
<tr><td></td><td>54</td><td></td></tr>
<tr><td>[h-]</td><td>42</td><td>牽 掀</td></tr>
<tr><td></td><td>31</td><td>賢 懸</td></tr>
<tr><td></td><td>35</td><td>犬 顯 憲</td></tr>
<tr><td></td><td>22</td><td>勸 獻</td></tr>
<tr><td></td><td>54</td><td>現</td></tr>
</table>

268

Zero 42 煙 冤 淵 豌 援
31 延 言 然 緣 元 圓 鉛 園 原 燃 丸 完 源 沿 筵 袁 員
35 軟 演 宴 腕 院
22 遠 燕 蔫 隕 焉
54 縣 願

(31) [-ip]

		左			右	

[p-] 42
31
35
22
54

[pʰ-] 42
31
35
22
54

[f-] 42
31
35
22
54

[m-] 42
31
35
22
54

[w-] 42
31
35
22
54

[t-] 42 帖
31
35
22
54

[tʰ-] 42 貼
31
35
22
54 迭 蝶 諜 疊

[n-] 42
31
35
22
54 聶 業

[l-] 42
31
35
22
54 獵

[ts-] 42 接
31
35
22
54

[tsʰ-] 42 妾
31
35
22
54 捷

[s-] 42
31
35
22
54

[j-] 42
31
35
22
54

[k-] 42 劫
31
35
22
54

[kʰ-] 42
31
35
22
54 俠 狹

[ŋ-] 42
31
35
22
54

269

[h-]	42	脅
	31	
	35	
	22	
	54	協 峽
Zero	42	
	31	
	35	
	22	
	54	葉

(32) [-it]

[p-]	42	別 必
	31	
	35	
	22	
	54	
[pʰ-]	42	
	31	
	35	
	22	
	54	
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	
	35	
	22	
	54	滅
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	
	31	
	35	
	22	
	54	跌 秩
[tʰ-]	42	鐵 脫
	31	
	35	

	22	
	54	奪
[n-]	42	
	31	
	35	
	22	
	54	
[l-]	42	
	31	
	35	
	22	
	54	列 劣 烈 裂
[ts-]	42	節 哲 □
	31	
	35	
	22	
	54	絕
[tsʰ-]	42	切 徹
	31	
	35	
	22	
	54	撤
[s-]	42	說 雪 設 吃 薛
	31	
	35	
	22	
	54	舌 蝕
[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	結 潔
	31	
	35	
	22	
	54	
[kʰ-]	42	缺 揭 決 竭 橛 蝎
	31	
	35	
	22	
	54	傑 □

[ŋ-]　42
　　31
　　35
　　22
　　54　熱
[h-]　42　血　闕
　　31
　　35
　　22
　　54　穴
Zero　42　乙
　　31
　　35
　　22
　　54　月　悅　越　粵

(33)　[-ik]
[p-]　42　碧
　　31
　　35
　　22
　　54
[pʰ-]　42
　　31
　　35
　　22
　　54
[f-]　42
　　31
　　35
　　22
　　54
[m-]　42
　　31
　　35
　　22
　　54
[w-]　42
　　31
　　35
　　22
　　54
[t-]　42　的
　　31
　　35

　　22
　　54
[tʰ-]　42
　　31
　　35
　　22
　　54
[n-]　42
　　31
　　35
　　22
　　54
[l-]　42
　　31
　　35
　　22
　　54
[ts-]　42
　　31
　　35
　　22
　　54　只　隻
[tsʰ-]　42
　　31
　　35
　　22
　　54
[s-]　42
　　31
　　35
　　22
　　54　食
[j-]　42
　　31
　　35
　　22
　　54
[k-]　42
　　31
　　35
　　22
　　54

Initial	Tone	Characters
[kʰ-]	42	
	31	
	35	
	22	
	54	
[ŋ-]	42	
	31	
	35	
	22	
	54	
[h-]	42	
	31	
	35	
	22	
	54	
Zero	42	
	31	
	35	
	22	
	54	

(34) [-iu]

Initial	Tone	Characters
[p-]	42	標
	31	
	35	表
	22	
	54	
[pʰ-]	42	彪 瓢 漂 飄
	31	嫖
	35	
	22	漂 票
	54	
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	苗 描
	35	秒 妙
	22	
	54	廟
[w-]	42	
	31	
	35	

Initial	Tone	Characters
	22	
	54	
[t-]	42	丟 刁 貂 雕
	31	
	35	
	22	釣 吊
	54	掉
[tʰ-]	42	挑
	31	條 調
	35	
	22	跳
	54	調
[n-]	42	
	31	
	35	
	22	
	54	尿
[l-]	42	
	31	寥 撩 遼
	35	了
	22	
	54	料 廖
[ts-]	42	招 焦 樵 澆 朝
	31	
	35	
	22	照
	54	
[tsʰ-]	42	超
	31	潮 朝
	35	
	22	
	54	兆 趙
[s-]	42	燒 悄 俏 簫 宵 消 硝 銷 霄
	31	紹 邵 韶 肇 召 沼 詔 昭
	35	小 少
	22	笑 少
	54	

[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	矯 嬌 驕
	31	
	35	繳 剿 □
	22	叫
	54	
[kʰ-]	42	
	31	橋
	35	
	22	
	54	轎 喬
[ŋ-]	42	
	31	
	35	
	22	
	54	
[h-]	42	
	31	
	35	
	22	
	54	
Zero	42	腰 妖 邀 要
	31	撓 擾 饒 曉 堯 姚 搖 謠
	35	
	22	繞 夭 要
	54	耀

(35) [-u]

[p-]	42	
	31	
	35	補 斧
	22	布 怖 鋪
	54	
[pʰ-]	42	
	31	菩 蒲
	35	普 譜 捕 甫 輔 浦
	22	鋪
	54	步 部 簿
[f-]	42	夫 乎 膚 俘 呼
	31	扶 符 胡 護 狐 壺 湖
	35	虎 府 腑 敷
	22	副 富 赴 褲
	54	付 婦 附 父 戶 俯 傅 腐 負 互
[m-]	42	
	31	無 模 巫 誣
	35	舞
	22	武 母 牡 畝
	54	募 墓 暮 務 霧
[w-]	42	
	31	
	35	
	22	
	54	芋
[t-]	42	都
	31	
	35	賭 堵 肚
	22	
	54	
[tʰ-]	42	途 屠 圖
	31	土 吐
	35	兔
	22	
	54	杜 度 渡 鍍
[n-]	42	
	31	奴
	35	努
	22	
	54	怒
[l-]	42	
	31	盧 驢 盧 爐
	35	鹵 魯
	22	
	54	路 露
[ts-]	42	租
	31	
	35	祖 組
	22	做
	54	

[tsʰ-] 42 粗
31
35
22 醋
54
[s-] 42 蘇 酥
31
35 數
22 數 漱 素 訴
54
[j-] 42
31
35
22
54
[k-] 42 姑 孤 辜 枯
31
35 古 苦 估 股 鼓
22 固 故 雇 顧
54
[kʰ-] 42
31
35
22
54
[ŋ-] 42
31
35
22
54
[h-] 42
31
35
22
54
Zero 42 污 烏
31
35
22
54

(36) [-ui]
[p-] 42 杯
31
35

22 貝 背 輩
54
[pʰ-] 42 丕 胚
31 賠 陪 培
35
22 倍 佩 沛
54
[f-] 42 開 恢 盃
31
35 海
22
54 害
[m-] 42
31 霉 枚 玫 梅 媒
35
22 每
54 妹
[w-] 42
31
35
22
54
[t-] 42 堆
31
35
22 對
54
[tʰ-] 42 推
31 檯 抬 臺 台
35 腿
22 褪 退
54 袋 隊 待 怠 代 貸
[n-] 42
31
35
22
54 內
[l-] 42
31 雷 來
35 縷
22 壘 屢
54 類 累 淚

[ts-]	42	哉 追
	31	
	35	宰 嘴
	22	醉 最 再
	54	
[tsʰ-]	42	吹 炊 崔 催
	31	才 錘 材 財 裁 隨
	35	采 彩 睬
	22	菜 趣 翠 賽
	54	罪 序 在
[s-]	42	雖 髓 須
	31	誰 垂
	35	水
	22	稅 帥 碎 歲
	54	瑞 睡 墅 遂 隧 穗 粹
[j-]	42	
	31	
	35	
	22	
	54	銳
[k-]	42	
	31	
	35	愧 改
	22	蓋
	54	
[kʰ-]	42	該
	31	
	35	概 凱 愷 慨
	22	
	54	
[ŋ-]	42	
	31	呆
	35	
	22	
	54	外
[h-]	42	哀
	31	回
	35	
	22	
	54	繪
Zero	42	
	31	
	35	

	22	愛 □
	54	會 為
(37) [-un]		
[p-]	42	搬 般
	31	
	35	本
	22	半
	54	
[pʰ-]	42	潘
	31	盤 盆
	35	
	22	判
	54	伴 胖
[f-]	42	寬 歡 鼾
	31	寒
	35	
	22	漢 旱
	54	汗
[m-]	42	
	31	門 瞞
	35	滿
	22	悶
	54	
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	
	31	
	35	
	22	
	54	
[tʰ-]	42	
	31	
	35	
	22	
	54	
[n-]	42	
	31	
	35	
	22	
	54	

[l-]	42	
	31	
	35	
	22	
	54	
[ts-]	42	
	31	
	35	
	22	
	54	
[tsʰ-]	42	
	31	
	35	
	22	
	54	
[s-]	42	
	31	
	35	
	22	
	54	
[j-]	42	
	31	
	35	
	22	
	54	
[k-]	42	竿 觀 官 乾 肝
	31	
	35	管 館 趕
	22	
	54	慣 冠 貫 灌 罐 幹
[kʰ-]	42	
	31	
	35	
	22	看
	54	
[ŋ-]	42	菫 喚 煥
	31	韓
	35	刊 罕 穎
	22	
	54	緩
[h-]	42	
	31	
	35	
	22	
	54	

Zero	42	安 鞍
	31	
	35	碗 宛 婉
	22	案 按
	54	換

(38)　[-uŋ]

[p-]	42	
	31	
	35	
	22	
	54	
[pʰ-]	42	
	31	馮
	35	
	22	
	54	
[f-]	42	風 蜂 峰 諷 瘋 豐
	31	
	35	
	22	
	54	鳳
[m-]	42	
	31	
	35	
	22	
	54	
[w-]	42	
	31	
	35	
	22	
	54	
[t-]	42	東 冬
	31	
	35	懂
	22	凍
	54	
[tʰ-]	42	通
	31	同 銅
	35	桶 統
	22	痛
	54	動 洞

276

[n-] 42
31 膿
35
22
54 弄

[l-] 42
31 龍 聾
35
22
54

[ts-] 42 中 忠
31
35 腫 種
22 粽 綜
54

[tsʰ-] 42
31 從 蟲 松
35
22
54

[s-] 42 雙
31
35 送
22 誦
54

[j-] 42 翁
31 融 容
35 涌
22
54 用

[k-] 42 公 宮 恭 功 工 攻
31
35
22 貢
54

[kʰ-] 42
31 窮
35
22
54 共

[ŋ-] 42
31
35

22
54

[h-] 42 胸
31 雄 紅
35
22
54

Zero 42
31
35
22
54

(39) [-ut]

[p-] 42
31
35
22
54 勃

[pʰ-] 42 撥
31
35
22
54 拔

[f-] 42 闊 忽 □
31
35
22
54

[m-] 42
31
35
22
54 没

[w-] 42
31
35
22
54

[t-] 42
31
35
22
54

277

[tʰ-]   42
        31
        35
        22
        54   突
[n-]   42
        31
        35
        22
        54
[l-]   42
        31
        35
        22
        54
[ts-]   42
        31
        35
        22
        54
[tsʰ-] 42
        31
        35
        22
        54
[s-]   42
        31
        35
        22
        54
[j-]   42
        31
        35
        22
        54
[k-]   42   割 葛
        31
        35
        22
        54
[kʰ-]   42
        31
        35
        22
        54
[ŋ-]   42

        31
        35
        22
        54
[h-]   42   歇 渴 喝
        31
        35
        22
        54   曷
Zero   42
        31
        35
        22
        54   活

(40) [-uk]
[p-]   42   撲
        31
        35
        22
        54
[pʰ-]   42
        31
        35
        22
        54
[f-]   42   福
        31
        35
        22
        54   服 復 覆
[m-]   42
        31
        35
        22
        54   木
[w-]   42
        31
        35
        22
        54

[t-]	42	督
	31	
	35	
	22	
	54	
[tʰ-]	42	
	31	
	35	
	22	
	54	毒 讀
[n-]	42	
	31	
	35	
	22	
	54	
[l-]	42	綠
	31	
	35	
	22	
	54	六 鹿 履
[ts-]	42	竹 足
	31	
	35	
	22	
	54	
[tsʰ-]	42	畜
	31	
	35	
	22	
	54	俗 族
[s-]	42	宿
	31	
	35	
	22	
	54	
[j-]	42	
	31	
	35	
	22	
	54	慾 肉 玉
[k-]	42	穀
	31	
	35	
	22	
	54	

[kʰ-]	42	曲 菊
	31	
	35	
	22	
	54	局
[ŋ-]	42	
	31	
	35	
	22	
	54	
[h-]	42	哭
	31	
	35	
	22	
	54	
Zero	42	
	31	
	35	
	22	
	54	

(41) [-m̩]

[p-]	42	
	31	
	35	
	22	
	54	
[pʰ-]	42	
	31	
	35	
	22	
	54	
[f-]	42	
	31	
	35	
	22	
	54	
[m-]	42	
	31	
	35	
	22	
	54	

[w-]	42			31	
	31			35	
	35			22	
	22			54	
	54		[kʰ-]	42	
[t-]	42			31	
	31			35	
	35			22	
	22			54	
	54		[ŋ-]	42	
[tʰ-]	42			31	
	31			35	
	35			22	
	22			54	
	54		[h-]	42	
[n-]	42			31	
	31			35	
	35			22	
	22			54	
	54		Zero	42	
[l-]	42			31	吳 吾 梧
	31			35	五 午 伍
	35			22	悟
	22			54	誤
	54				
[ts-]	42				
	31				
	35				
	22				
	54				
[tsʰ-]	42				
	31				
	35				
	22				
	54				
[s-]	42				
	31				
	35				
	22				
	54				
[j-]	42				
	31				
	35				
	22				
	54				
[k-]	42				

Appendix D: Capitalized Abbreviations

AD adverb

ADJ adjective

ADV adverbial phrase

ASP aspect

AUX auxiliary

C consonant

CL classifier

CRS currently relevant state

DEM demonstrative

DET determiner

DIM diminutive

DP determiner phrase

EX expected (to be confirmed

positively)

EXP experiential aspect

F final endpoint

IMP imperfective aspect

INC inchoative

M medial

NCL numeral classifier

NEG negative

NOM nominal

NP noun phrase

NUM numeral

O syntactic object

PASS passive particle

PFV perfective marker

PL plural

POS possessive

POT potential

PP prepositional phrase

PREP preposition

PRO pronoun

PROG progressive aspect

PRT particle

Q question particle

REL relative marker

SA solicit agreement

SG singular

V vowel or verb, depending on context

VP verb phrase