THE UNIVERSITY *of* EDINBURGH

Unpacking the History of Middle Chinese *ɣu- in the Yue Dialects in Guangdong: a Dialect Geographical Analysis

THE UNIVERSITY *of* EDINBURGH

Name: Ho Wang Matthew Sung

Word count: 32864

Name of Degree: MSc Linguistics by Research

The University of Edinburgh

2020

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgment, the work presented is entirely my own.

Abstract

The main focus of this dissertation is the developments of the Middle Chinese sequence *ɣu- in 54 Yue dialects in Guangdong. The reflexes of MC *ɣu- sequence invites attention because (a) the traditional dialect description does not capture the developments of this sequence across all the dialects systematically, (b) the phonological history of *ɣu- in Yue remains in the descriptive stage without reference to sound changes involved between Middle Chinese and present-day Yue dialects and (c) the inter- and intra-dialect variation of the reflexes of *ɣu- is still unexplored.

I use Yue dialect survey data (Zhan & Cheung 1987, Zhan & Cheung 1994, Zhan & Cheung 1998, Shao 2016 and Beijing University Linguistics Faculty 1989) to answer the following research questions:

1) Are there geographical patterns for the variation of the reflexes of MC *ɣu- in Yue?
2) Do all the words have the same MC *ɣu- reflex within one dialect?
3) How many steps were involved in the changes from Middle Chinese to present-day varieties?
4) Were there specific phonological contexts for certain changes to occur?
5) Was there contact or diffusion of these changes between Yue and other neighbouring varieties such as Hakka?
6) Did linguistically similar but geographically distant patterns develop from the same changes?

The data shows that based on the reflexes of MC *ɣu-, (1) Yue dialects in Guangdong can be divided into four groups, (2) there can be more than one reflex for MC *ɣu- within a dialect, (3) different dialect groups have different number of steps from MC *ɣu- to the present-day reflex, (4) most dialects show that MC *-u- triggers ɣ-loss in the history of Yue, (5) I propose that the current geolinguistic pattern resulting from the contact with Hakka is unlikely; the current dialect landscape was created by contact between Yue varieties, and lastly (6) two groups of dialects (Bao'an- and Maoming-type dialects) appear to be linked historically despite them being separated by the Taishan-type dialect geographically.

Acknowledgements

I would like to thank my supervisors, Patrick and Warren. They have always been supportive and kind to me and supported me throughout the year. Patrick has not only been supervising my dissertation, he has also helped me a lot with other things such as my PhD application and the Carnegie scholarship application. Other than supervision, Warren has been helping me with dialectology, e.g. by lending me books I needed for a better understanding of dialectometry and answering my questions on various topics in dialectology.

I would also like to thank the Ailie Donald Scholarship for the financial support this year. This scholarship has reduced the financial burden for me and my family.

There is another person I have to mention here, Pavel. He is the person who pointed me to the direction of applying western dialectological methods to the Chinese data, which seemed obvious, but nevertheless changed my way of thinking since. Without him, my dissertation and research interests would be very different.

This dissertation has applied techniques which I received help and guidance from outside the university. The GIS mapping technique comes from the help of Dr. Chris Montgomery from Sheffield. For dialectometry and the use of Gabmap, the guidance and help came from Dr. Wilbert Heeringa, Prof. John Nerbonne and Prof. Jelena Prokic from the Netherlands.

I want to thank Nina for her help. Despite how busy she was, she was willing to look at a section of my dissertation and gave me some really useful feedback.

To get hold of the data and the reading material for Chinese dialectology and historical phonology, I could not have done it without the help of my friends who helped me to get all the books I needed. They are Ziming Lu, Masa Guan, Yingying Pang, Kate Chen, Zhenshen Zhou.

Lastly, I'd like to thank my parents for letting me follow my passion. Without them, I would not be able to pursue this route and I cannot thank them enough for this.

Contents

List of Tables

List of Figures

1. Introduction

Yue is a group of dialects spoken mainly in the Guangdong and Guangxi provinces in southern China. Cantonese, which is spoken by a large population in Hong Kong and Guangzhou, belongs to Yue (Bauer & Benedict 1997: xxxi). Amongst the Sinitic languages, Yue is renowned for the retention of some Middle Chinese (MC) features, namely the preservation of all MC nasal and stop codas. However, Yue has certainly changed since the Middle Chinese period, as would be expected. One of the changes concerns the developments of the MC sequence *ɣu-. In (1), I have listed two words and given their MC reconstruction and their pronunciation in four present-day Yue dialects from different parts of Guangdong. We can easily spot complicated dialect patterns. For instance, Guangzhou [w-] corresponds to [v-] in Taishan, but that is not necessarily the case for Dongguan, as Guangzhou [w-] corresponds to both [f-] and [v-] in Dongguan. Similarly with Maoming, except the initial consonantal correspondence is a [ʋ-] instead of a [v-].

(1)

	Middle Chinese[1]	Guangzhou	Taishan	Dongguan	Maoming
'to return'	*ɣuan	[wan]	[van]	[van]	[ʋan]
'lake'	*ɣo/*ɣu	[wu]	[vu]	[fu]	[fu]

The exact reasons behind such inter- and intra-dialect variation across Guangdong have not been explored in detail before. The phonological changes involved in the examples above have also not been systematically described. This is partly due to the methodology in traditional Chinese dialectology. The standard procedure in the documentation of Chinese dialects (phonetics and phonology) involves transcribing around 1000 to 3000 characters and publishing a report which compiles the sound correspondences of Middle Chinese rhyme book categories and their reflexes in the dialect. The shortcoming of this approach is that we often

[1] There are many Middle Chinese (MC) reconstructions (e.g. Tung 1968, Li 1973, Wang 1985, Zhu 2016). This dissertation uses Zhu's (2016) reconstruction. See Section 3. There are two MC reconstructions of 'lake', this is because in different periods of MC there are different realisations. See Section 5.2.3 for the explanation.

do not know which precise changes were involved between Middle Chinese and present-day dialects since the analysis is mostly descriptive and the changes involved are open to interpretation. To move a step forward from dialect description, it is necessarily to look deeper into why such dialectal variation exists and how it relates to the historical changes with MC *ɣ-, which is absent in all present-day dialects.

This dissertation focuses on the reflexes of the MC sequence *ɣu- in present-day Yue dialects in Guangdong. The concept of a sequence derives from Chinese dialectology. The traditional terminology for describing the set of words which had a *ɣ- initial before a medial *-u- in Middle Chinese, such as the ones in (1), is *匣母合口字* [words with MC *ɣu-]. The traditional description highlights both the segment which requires attention and its phonological context. This dissertation follows that tradition.

My focus is specifically on the *ɣu- sequence and not simply on the segment *ɣ- because MC *ɣ- went through different changes in different phonological environments. MC *ɣ- before MC *-u- shows a huge amount of inter- and intra-dialectal variation, such as that shown in (1) whereas MC *ɣ- became devoiced and debuccalised to [h-] in present-day dialects without much dialect variation in the elsewhere environment. In the literature, there is a lack of systematic description of the development of MC *ɣu- across dialects of Yue. Hence, the variation of the reflexes of *ɣu- is currently unexplored. Furthermore, existing dialect descriptions only generalise about the changes that affected MC *ɣ-, without making reference to its phonological environments in detail (e.g. Zhan 2002). Other studies (e.g. Chen & Newman 1984a, Li 2015) are rather dialect-specific, which result in a very partial and fragmented picture of the developments of *ɣu- in the Yue-speaking area in Guangdong.

Given all this, I aim in this dissertation to apply an alternative methodology to investigate the

sound changes that affected this sequence between Middle Chinese and present-day Yue dialects. By examining the present-day reflexes of the MC sequence *ɣu-, I move beyond the descriptive stage in Chinese dialectology to explore the developments between Middle Chinese and present-day Yue dialects. To do this, I make use of a mixed approach, which combines aspects of both Western dialectology and traditional Chinese historical phonology. This approach does not only allow us to explore the historical changes further, it also makes Chinese dialectology and historical phonology more accessible for scholars who are not Chinese readers or Sinologists.

My research questions of this dissertation are as follows:

1. Are there geographical patterns for the variation of the reflexes of MC *ɣu- in Yue?
2. Do all the words have the same MC *ɣu- reflex within one dialect?
3. How many steps were involved in the changes from Middle Chinese to present-day varieties?
4. Were there specific phonological contexts for certain changes to occur?
5. Was there contact or diffusion of these changes between Yue and other neighbouring varieties such as Hakka?
6. Did linguistically similar but geographically distant patterns develop from the same changes?

To address these questions, I analyse the Chinese dialect survey data with modern dialectological methods. Data are extracted from several dialect surveys which were conducted from the late 1980s until the end of the 1990s, and one from the 2010s. The dialect survey data were mainly collected from local, non-mobile speakers who were proficient in their own local dialects. I use several contemporary dialectological methods in this dissertation, including multi-dimensional scaling (MDS) and point-colour maps. MDS is useful for showing a more

objective relationship between dialects and point-colour maps are useful to show the geolinguistic patterns of a variant/reflex.

This dissertation is structured as follows: Section 2 gives a brief introduction to Guangdong province, and to the geography, history and languages of Guangdong. This section offers the background needed to understand the formation and the geographical proximity of the dialects in Guangdong. Traditional terminologies usually form a barrier which makes Chinese dialectology not accessible to scholars who do not read Chinese. Section 3 introduces the methodology and terminology in traditional Chinese dialectology and the definition of the term 'dialect' used in the dissertation. Section 4 gives a description of what Yue is, the formation of Yue and the classification of dialects within Yue. Next, the MC sequence *ɣu- is explained in detail in Section 5. This section introduces previous research on the *ɣu- sequence in Yue dialects, which also provides an important basis for the analysis in this dissertation. Section 6 gives an overview of the data used in this dissertation. Section 7 comprises the main analysis. This section introduces the application of 'western' dialectological methods and presents the findings. The discussion is in Section 8. Finally, the whole dissertation is summed up in Section 9.

2. Guangdong Province

2.1 Geography of Guangdong

Guangdong, also known as Yue in short, is one of the most southerly provinces in China, situated in the South-east, facing the South China Sea. To its West lies the Guangxi province; Fujian is to its East and lastly Hunan and Jiangxi are to its North. In addition, Hainan Island lies to the south of the western part of Guangdong. There are also two Special Administrative Regions (SARs) to the south of Guangdong, namely Hong Kong and Macau. The location of Guangdong in China is shown in Figure 1.

Figure 1. Guangdong and two SARs in relation to China and Taiwan[2]

The size of the province is 178000 km^2. The coastline of Guangdong is 3368.5 km, which is the longest of any province in China (Wu & Zhan 2008: 109). There are several types of landscape in the province: a) Mountainous areas, b) plateaus, c) hilly areas, d) terraces and e) islands (Wu & Zhan 2008: 109). Mountainous and hilly areas cover 60% of the province,

[2] This map and the rest of the maps in this dissertation are made with *QGIS* (QGIS Development Team 2020); the base map (shapefile) was downloaded from https://maps.princeton.edu/catalog/harvard-chgis-v4-1997-prov-pgn

terraces cover 16.5%, plateaus cover 23.7% and there are 759 islands. There are 9 rivers in total, with the Xijiang, Beijiang and Dongjiang being the major rivers which form the Pearl River river system - one of the four major river systems in China. A map of Guangdong with physical geographical information can be found in Figure 2. Red lines indicate major roads.

Figure 2. Physical geographical map of Guangdong (Wiegand 2007: 68-69)

2.2 A brief history of Guangdong

Before the Qin dynasty (221 – 207 BCE), the Hua-Xia people (roughly corresponds to 'Chinese' nowadays) lived around the Yellow River, known as the Central Plateau area. Present-day Guangdong was seen as the land of the barbarians in ancient times (Zhan 2002: 2); it belonged to the Nanyue area then, which was part of the wider Baiyue area. Baiyue literally means 'Hundred Yues' (Yuan 2001: 177, Chinagate 2009). It should be noted that the people living in this area were not Han people (also corresponding to 'Chinese' nowadays) at the time.

The first Emperor of the Qin dynasty unified southern China (Lingnan, which includes present-day Guangdong) in 214 BCE and the Nanhai Commandery was established there (Wu & Zhan 2008: 109). After the death of the first Emperor of Qin, Zhaotuo, one of the generals who led the conquest of Lingnan, seized three counties in Lingnan and called it the Southern Yue Kingdom, declaring independence from the Qin dynasty. The capital of Zhaotuo's Kingdom

was located in Panyu, which is present-day Guangzhou. At the beginning of the West Han dynasty (202 BCE – 9 CE), the Emperor of the Han dynasty regained this territory and renamed it as the Nanhai Commandery again in 111 BCE (Chinagate 2009, Wu & Zhan 2008: 109).

In 226 CE (during the Three Kingdom Period, 220-280 CE), the present-day Guangdong area was named the Guang province. It stayed the same during the Jin (266-420) and the Northern and Southern Dynasties (420-589 CE). The Wen Emperor of the Sui dynasty (581-619 CE) abolished the Nanhai Commandery, and it was not reestablished until 607 CE. In 621 CE, the Tang dynasty (618-907 CE) reestablish the Guang province. In 758 CE, a regional commander from the Tang dynasty was put in charge of present-day Guangdong. In 862 CE, Lingnan was split into East and West; present-day Guangdong belongs to the Eastern part (Chinagate 2009).

At the end of the Tang dynasty, the present-day Guangdong area became the centre of another regime – the Southern Han Kingdom. This was the case until the first emperor of the Song dynasty (960-1279 CE) abolished the Southern Han Kingdom and restored Lingnan in 971 CE. In 997 CE, the Song dynasty established the Guangnan East Circuit, which is where the name Guangdong (literal meaning: Guang East) comes from (Chinagate 2009).

During the Ming dynasty (1368-1644 CE), present-day Guangdong was named as the Guangdong province in 1376 CE. In 1911, the Guangdong province declared independence and became a jurisdiction under the Republic of China. From 1938 – 1945, some areas of Guangdong were occupied by the Japanese army (Chinagate 2009).

After the establishment of the People's Republic of China (1949), the political region of Guangdong has been modified. For example, Huaji was moved from Guangxi to Guangdong; Qinzhou, Fangchenggang and Beihai became part of Guangxi etc. (Chinagate 2009).

2.3 Languages and dialects in Guangdong

In Guangdong, people mainly speak Sinitic varieties, but there are also speakers of minority languages such as Zhuang, Yao (Mien), and She (Wu & Zhan 2008: 110).

According to Wu & Zhan (2008: 110), the main Sinitic varieties spoken in Guangdong are Yue, Hakka and Min. In addition, Yuebei Tuhua (literal translation: Northern Guangdong Local Vernacular) and Junhua (literally means 'Army language'; it is related to Mandarin/Guanhua), are also spoken in the province. In some areas, people speak Southwestern Mandarin. See Figure 3 for the distribution of speakers of these dialects. There are more discussion of Yue dialects in Section 4.

Figure 3 is taken from the English version of the *Language Atlas of China* (Wurm et al. 1989: B13). This map shows the dialect areas of a number of Sinitic varieties spoken in Guangdong. Note that a second edition of this atlas was published in 2012. I use the older edition because: 1) the new edition is only in Chinese and 2) the cartographic techniques used in the new edition is hard to interpret. Therefore, I believe this map shows a clearer picture of the dialect landscape of Guangdong for all readers.

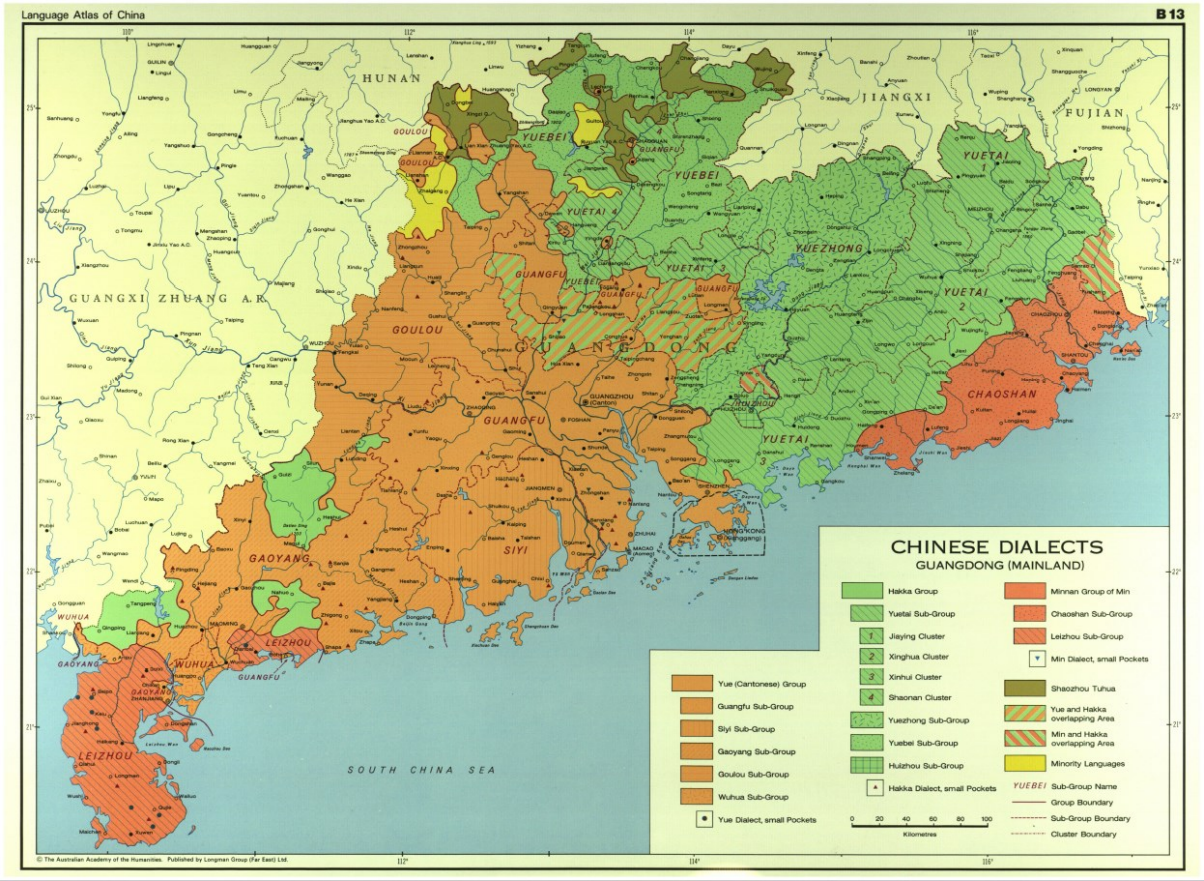

Figure 3. Linguistic map of Guangdong (Wurm et al. 1989: B13)

We can see in Figure 3 that traditional dialects of Yue were mainly spoken in the western half of the province (in orange), whilst Hakka was spoken in the eastern half (in green), with some sprachinseln near Guangxi. There were also bilingual areas of Yue and Hakka in central Guangdong. Traditional Min varieties are spoken on the edges of Guangdong (in red), namely in the Chaoshan area in the East and on the Leizhou Peninsula in the West. Lastly, there are Tuhua varieties (dark green) and other minority languages (yellow) spoken in the north.

3. Traditional Chinese Dialectology and Historical Phonology

Traditional Chinese dialectology and historical phonology are inseparable. From data collection to dialect classification, one can hardly avoid references to Middle Chinese rhyme book categories. In this section, the framework, methodology and terminologies from traditional Chinese dialectology and historical phonology are introduced. It is very important to gain an understanding of these traditional terminologies because they are widely used in the literature and readers are often assumed to know what they mean. Readers are warned that only the most important terminologies needed for Chinese dialectology and terminologies related to the focus of this dissertation are introduced; other details can be found in Zhu (2016).

This dissertation uses the reconstructed values for the MC sound categories in order to depart from simple correspondence description between the reflexes and the MC categories. With the reconstructed value, it becomes immediately clearer what changes could be involved between MC and present-day Yue dialects. Unless indicated, this dissertation uses Zhu's (2016) reconstruction because this reconstruction is based on a comparison with several reconstructions of MC (e.g. Tung 1968, Li 1973, Wang 1985) and Zhu assessed historical and present-day dialectal evidence before offering his reconstruction. The current dissertation is not to judge whether this reconstruction is correct or not. I provide Zhu's (2016) reconstruction simply to give a historical reference point for readers to refer to and also to analyse phonological changes. Some alternative reconstructions are given (see Section 5.2.3, 8.2) in order to account for different analyses and potential historical dialectal differences.

3.1 Traditional Chinese Dialectology

3.1.1 Framework

Chinese dialectology works with the assumption that present-day dialects (except Min) are descended from the phonological system of Middle Chinese, represented by the rhyme book

Guangyun[3] (You 2016: 86). The phonological system of *Guangyun* is "an ideal phonological system for comparing dialect data" (You 2016: 85), since the sound categories listed in *Guangyun* capture all the contrasts present in modern dialects and dialect features can be identified through looking at the correspondences between the dialect and *Guangyun* (You 2016: 85).

3.1.2 Methodology

The *Dialect Survey Wordlist* (*Wordlist* hereafter, Chinese Academy of Social Science (CASS) 1983) is used by Chinese dialectologists for data collection. It is organized according to the sound categories in the MC rhyme book *Guangyun*, for ease of identifying sound correspondences to Middle Chinese (You 2016: 55). 'Word' here refers to mono-syllabic words. Poly-syllabic words are not considered in this wordlist because in *Guangyun*, only mono-syllables were recorded.

In Chinse dialectology, the direct method (Chambers & Trudgill 1998: 21) is used for data collection. Fieldworkers go into the community to conduct their fieldwork (You 2016: 54-55). Informants are asked to read out the characters from the *Wordlist*. This method is relatively easy to conduct and it can yield a lot of data in the process in a short period. Lastly, the pronunciation of these characters are transcribed in a modified IPA[4]. Section 6 expands on the informant selection process.

In dialect description, sometimes only a small number of correspondences between Middle Chinese sound categories and reflexes of present-day dialect are described. There can be no mention of the reconstruction of the sound value of these categories at all. It requires readers

[3] See Section 3.2
[4] See Section 6.3.

to have some knowledge of the Middle Chinese rhyme book and rhyme table categories[5] in order to understand the literature.

3.1.3 Data presentation

The data collected by the fieldworkers are usually published in several forms: 1) dialect survey reports (including the data), 2) monographs, 3) dialect dictionaries and 4) dialect maps/atlases (Li 2017: 162-184). This dissertation uses data from the data from dialect surveys. See Section 6 for more details about the data collection and data presentation of the Yue dialects surveys.

3.2 Chinese Historical Phonology

In Section 3.1, I have explained that traditional Chinese dialectology is closely related to traditional historical Chinese phonology, namely in the use of Middle Chinese sound categories. This section gives more explanation of the terminologies Chinese dialectology refers to.

3.2.1 *Qieyun*, *Guangyun*, *Fanqie* and rhyme tables

The following subsections introduce rhyme books (*Qieyun*, *Guangyun*); *Fanqie*, which is a way to analyse syllables; rhyme tables as well as sound categories for initials, finals (medials and rhyme) and tones.

3.2.1.1 *Qieyun* and *Guangyun*

Guangyun is a rhyme book compiled around 1007-1008 in the Song dynasty (960-1279 CE). The full name of *Guangyun* is "*Great Song revised and expanded rhymes*". It is a later edited version of a rhyme book of *Qieyun* (Sun 2018: 35), compiled by Lu Fayan and his colleagues in 601 CE in the Sui dynasty (581-619 CE). The original full copy of *Qieyun* is lost, but fragments have been found (Sun 2018: 32). These fragments preserve the introduction of *Qieyun* and the phonological descriptions of some characters. *Guangyun* is also known as a *Qieyun*-type rhyme book, because it reflects the phonological system of *Qieyun*, despite it

[5] Explained below in Section 3.2.

being compiled several centuries later (Zhu 2016: 184). This implies that the framework in Chinese dialectology assumes present-day dialects to descend from the *Qieyun* phonological system as well (see Section 3.1.1).

The exact phonological system that *Qieyun/Guangyun* reflects should be considered cautiously. There are different theories proposed regarding to the phonological system it represents (Chang 2019: 222): 1) the phonological system of the Chang'an (the capital of the Sui dynasty) dialect in the 6th century (Karlgren 1949), 2) a combination of two phonological systems (north and south) in the 6th century (Zhou 1966) and 3) two phonological systems (north and south) from different periods (no information of exactly when, Wang 1985). No matter what variety *Qieyun* actually represents, Chinese dialectology still bases the assumption that *Qieyun* is the ancestor of most present-day dialects in its methodology.

3.2.1.2 *Fanqie*

Fanqie is a method of analyzing syllables in ancient China, found in *Qieyun* and *Guangyun*. In western phonology, the syllable template consists of an onset, nucleus and coda. The nucleus and coda together form a rhyme. *Fanqie* splits a syllable into two parts, namely an initial, which is the onset of the syllable, and a final, which consists of everything else in the syllable, including the tone. This is demonstrated in Table 1.

		Character	Represented Initial	Represented Final	Represented Tone
Fanqie	Former character	作	*ts	*ɑk	*Ru*
	Latter character	孔	*k	*uŋ	*Shang*
Represented syllable		總	*ts	*uŋ	*Shang*
			*tsuŋ (with Shang tone)		

Table 1. Demonstration of *Fanqie* (adapted and modified from Sun 2018: 16)

Table 1 shows how the syllable 總 *tsuŋ is analysed in *Fanqie*. Firstly, the initial and final are represented by two characters, 作 *tsɑk and 孔 *kuŋ. The former character represents the initial and the latter represents the final of the syllable in question. To find the pronunciation, we take the initial from the former character and the final and tone in the latter, then merge them together. The dichotomy of an initial and a final forms the basis for syllable analysis in Chinese dialectology and historical phonology.

3.2.1.3 Rhyme tables

A rhyme table is a syllabary in short. The earliest rhyme tables (e.g. *Yunjing* and *Qiyunlüe*) were published in 1161 and 1160 CE respectively during the Song dynasty (c.f. *Qieyun* was published in 601CE). These early rhyme tables are seen as complementary to *Qieyun* for the reconstruction of Early Middle Chinese (Zhu 2016: 254).

In a rhyme table, characters were placed in different slots in a rhyme table to show its pronunciation. Each slot and table illustrate categories, such as *initials*, *rhyme group*, *vowel grade*, *medials* & *open/closed* rhyme categories and tones. These categories are very important and useful in Chinese dialectology as they are used as reference points in present-day dialect description (You 2016: 88). These categories are explained below.

It should be mentioned that the concept of a *rhyme group* was introduced several hundred years after *Qieyun*. It has been brought up by Hill (2019: 95) that using these later rhyme table categories to look at *Qieyun* categories shows "distort[ion to] one's perception of the *Qieyun*", that this framework accepts "anachronism of using a book from 1161 to analyse the categories of a book from 601".

This dissertation follows the sound categories used in Chinese dialectology and traditional

historical phonology as reference points to identify words that descend from MC *ɣu-. The anachronism problem is not addressed in this dissertation.

3.2.2 Initials

Initials in Middle Chinese are represented by characters like the one found in Table 2. Each character represents one sound category, which is usually reconstructed as a consonant, except for the *Yi* initial , which is reconstructed as a zero-initial *Ø (see Table 2). In Western phonology, consonants are described with three parameters: place of articulation, manner of articulation and voicing. In Chinese historical phonology, place and manner of articulation for initials are fused together, along with abstract categories such as 'clear' vs. 'muddy'.

'Clear' and 'muddy' correspond to 'voiceless' and 'voiced'. There is also two subtypes of 'clear' and 'muddy' – 'full' and 'secondary'. It has been agreed by many scholars that 'full clear' means 'voiceless unaspirated' and 'secondary clear' is 'voiceless aspirated' (e.g. Tung 1968, Li 1973, Wang 1985). For 'muddy' sounds, 'full' refer to 'voiced' and 'secondary' is 'sonorant', e.g. nasals and laterals. Table 2 gives a list of initial terminologies used in Chinese dialectology with reconstruction of their values in Middle Chinese (by Zhu 2016).

In Table 2, the place and manner of articulations are shown in the column on the left. The 'muddy' vs. 'clear' categories as well as 'full' and 'secondary' are simplified to distinctive features for ease of reading.

	-voi, -s.g.	-voi, +s.g.	+voi	+nas	+lat
Bilabial stops and nasal	幫 p	滂 p^h	並 b	明 m	
Labio-dental affricates and nasal	非 pf	敷 pf^h	奉 bv	微 ɱ	
Alveolar stops and sonorants	端 t	透 t^h	定 d	泥 n	來 l
Alveolar affricates	精 ts	清 ts^h	從 dz		
Alveolar fricatives	心 s		邪 z		
Postalveolar / retroflex affricates	莊 $\widehat{tʃ}$/$\widehat{tʂ}$	初 $\widehat{tʃ^h}$/$\widehat{tʂ^h}$	崇 $\widehat{dʒ}$/$\widehat{dʐ}$		
Postalveolar / retroflex fricative	生 ʃ/ʂ				
Alveolo-palatal affricates	章 tɕ	昌 $tɕ^h$	船 dʑ		
Alveolo-palatal fricatives	書 ɕ		禪 ʑ		
Palatal/ retroflex stops and nasal	知 c/ʈ	徹 c^h/$ʈ^h$	澄 ɟ/ɖ	娘 ɳ	
Velar stops and nasal	見 k	溪 k^h	群 g	疑 ŋ	
Velar Fricatives*	曉 x/h		**匣 ɣ/ɦ**		
Miscellaneous	云 $ɣ^j$	以 Ø			
Glottal stop	影 ʔ				
Nasal affricate	日 nz̩				

Table 2. Traditional layout of the Middle Chinese initials used in Chinese dialectology (adapted from You 2016: 87 with Zhu's 2016 reconstruction)

Note that for some initials, there are two reconstructions because those initials have a controversial status regarding their reconstructed values, e.g. the retroflex stops and affricates. These competing reconstructions are shown in Table 2 for reference. Because most of them are not relevant to this current study, I have omitted the details of the debates and their explanations. See Zhu (2016: 301-328) for more.

The *Xia* initial, *ɣ- in Table 2, requires a more detailed discussion, as it is the focus of this dissertation. An alternative reconstruction of the *Xia* initial is *ɦ-. Karlgren has provided arguments for both reconstructions (cited from Zhu 2016: 322-323). The arguments for the *ɣ- reconstruction is based on the evidence from Chinese loanwords/transcriptions in other Sino-Xenic languages, that the corresponding consonant is velar. In present-day dialects, the reflex in Mandarin is [x-] and in southern dialects, [k-] and [k^h-] in some words. They serve as the

arguments for reconstructing the initial as a velar consonant.

An argument for *ɦ- comes from the present-day dialects as well. The Xia initial is lost in many dialects; this is explained with ɦ-loss. H-dropping has been attested in English, Swedish and some Romance languages, and Karlgren believed ɦ- in theory should be even easier to be lost, which explains the zero reflex in many present-day dialects (Zhu 2016: 323). Moreover, the rhyme tables show that this initial is grouped in the glottal category instead of velar. In addition, words with the *Xia* initial were used to transliterate Sanskrit syllables with onset vowels. Karlgrens believed that this is because perceptually ɦ-headed and an onsetless syllables sound very similar.

Karlgren's view is that *ɣ- was an earlier form and *ɦ- was a later development. When *Qieyun* was compiled, Karlgren believed that in some southern dialects, *ɣ- > *ɦ- occurred (Zhu 2016: 323). Hence, different reflexes are found in present-day dialects. This dissertation uses the *ɣ- reconstruction.

Finally, initials are written with a hyphen following the consonant, e.g. *ɣ-.

3.2.3 Medials

A medial is the first element of a final (if present), preceding the nucleus of a syllable (Zhu 2016: 30). An important rhyme table category for medials, known as *Hu*, is still used in dialect descriptions today (You 2016: 91). It is essential to know what these categories mean in order to extract the data from the dialect surveys.

In the Middle Chinese rhyme tables, there are only two *Hu* categories: *hekou* and *kaikou* categories (You 2016: 91). Examples from present-day Guangzhou dialect (a Yue variety) and Middle Chinese are given for these two categories in Table 3.

Middle Chinese *Hu*	English name (e.g. in Chan 2006:44)	Sample word	Middle Chinese	Guangzhou Yue
Kaikou	*Open rhyme*	蝦 'prawn'	*ɣa	[ha]
Hekou	*Closed rhyme*	華 'Chinese'	*ɣua	[wa]

Table 3 *Kaikou and Hekou* categories with examples in Middle Chinese and Guangzhou Yue

As Table 3 shows, the Middle Chinese word in the *kaikou* category has no medial *-u-, whereas in the *hekou* category, a medial *-u- is present, as shown in the word 'Chinese'. It should be noted that rhymes such as *-u (*Yu* rhyme group[6]) or *-uŋ (*Tong* rhyme group) are considered as a *hekou* rhyme[7] as well. For the ease of the readers, I am using the English translations *Open* and *Closed* rhymes for *kaikou* and *hekou* (e.g. in Chan 2006: 44) in the rest of the dissertation. Lastly, since the focus of this dissertation is *ɣu-, only the *hekou*/*closed rhyme* category is relevant.

In Chinese dialectology, there are many cases where an initial is analysed together with the medial. For example, Zhan (2002: 109) stated that the MC *Xiao* initial in combination with the open rhyme (*曉母合口*, *xu-) is reflected as [f-]. In this statement, the initial is not analysed as one segment, but as a sequence. A sequence like this tells us that the reflex does not only correspond to the initial, but also the medial. It also states the phonological context for any changes that occurred to the initial. I follow this tradition in this dissertation.

Lastly, medials are represented by a high vowel between two hyphens, e.g. *-u-.

3.2.4 Rhymes and rhyme groups

A rhyme is everything within the final other than the medial (i.e. nucleus and coda, Zhu 2016: 176). In Chinese dialectology, the rhyme is analysed as a unit. Vowel phonemes are not analysed separately and codas are sometimes analysed independently with reference to the

[6] See Section 3.2.4 for the explanation of rhyme groups.
[7] Based on Zhu (2016: 340-341, 351) reconstruction.

relevant rhyme groups. This aspect appears to be a big difference to western phonology.

While it is possible to talk about individual rhyme categories found in *Qieyun*-*Guangyun*, dialectologists usually describe the rhymes in a dialect with reference to the Middle Chinese rhyme groups[8] (You 2016: 88). Chinese dialectologists uses the sixteen rhyme groups categorized by ancient Chinese philologists to describe present-day dialects (You 2016: 88-89). These sixteen rhyme groups act as overarching categories for 61 different rhyme categories in *Guangyun*. These rhyme groups are listed in Table 4 and the explanation is provided below.

Name (English)	*Does it have closed rhymes*?	*Ru* category?	Vowel grades:	Can *ɣu- combine with this rhyme group (Grade I and II)?
Guo	yes	no	I, III	yes
Jia	yes	no	II, III	yes
Yu	yes (no *open rhymes*)	no	I, III	yes
Xie	yes	no	I, II, III IV	yes
Zhi	yes	no	III	no
Xiao	no	no	I, II, III, IV	no
Liu	no	no	I, III	no
Xian	yes	yes	I, II, III, IV	no
Shen	yes	yes	I, II, III, IV	no
Shan	yes	yes	I, II, III, IV	yes
Zhen	yes	yes	I, III	yes
Dang	yes	yes	I, III	yes
Jiang	no	yes	II	no
Zeng	yes	yes	I, III	yes
Geng	yes	yes	II, III, IV	yes
Tong	yes (no *open rhymes*)	yes	I, III	yes

Table 4. List of rhyme groups

Different Middle Chinese rhymes were assigned to a rhyme group based on the similarity of their nucleus and codas (You 2016: 88). An example taken is given in Table 5:

[8] This is the reason why individual rhyme categories of *Guangyun* are not introduced in this section.

Vowel Grades	*Open rhymes*	*Closed rhymes*
I	-ɑn, -ɑt	-uɑn, -uɑt
II	-an, -at	-uan, -uat
III	-jæn, -jæt	-juæn, -juæt
IV	-ien, -iet	-iuen, -iuet

Table 5. Rhymes in the *Shan* rhyme group (adapted from Zhu 2016: 349)

Table 5 consists of reconstruction of the rhymes within the *Shan* rhyme group[9]. The *Shan* rhyme group is an excellent example for explaining rhyme table categories. The rhymes in Table 5 were grouped together because they share a non-high and mostly front vowel with a coronal nasal/stop coda. Next, the rhymes are split into the *open* and *closed* rhyme categories within a rhyme group (see Section 3.2.3 for the explanation). The vowel grades are to distinguish the nucleus of the rhyme by their vowel height (Zhu 2016: 244, You 2016: 91). The MC Grade I vowel tends to be low and back and the Grade II, III and IV vowels are front with different vowel heights (II is the lowest and IV is the highest, see Table 5). Furthermore, Grade III and IV indicate the presence of the medial *-j- and *-i- respectively (Zhu 2016: 336). For Grade III or IV closed rhymes, two medials coexist in the same rhyme, as shown in Table 5.

Back to Table 4, it provides some details relevant to this current dissertation. Firstly, not all rhyme groups has a closed rhyme (see *Tong* and *Xie* rhyme groups), i.e. not having a *-u-. Next, the *Ru* category here refers to whether the rhyme group has rhymes with stop codas or not. These stop codas are only found in rhyme groups which have a nasal coda (homorganic to the stops), see Table 5 for an example. This is an economic way to group rhymes which differ in the coda only. Lastly, the last column of Table 4 introduces the rhymes that are of interest for this dissertation. The scope of this dissertation only focuses on the MC *ɣu- sequence in

[9] The reconstruction of all rhyme groups is not given because there are too many different versions in the literature and the main concern of this dissertation is not on Middle Chinese rhymes.

combination with Grade I and II rhymes. With Grade III and IV rhymes, there could be potential interactions between two medials (*-u- and *-i-/ -j-). Before understanding the sequence with only one medial (i.e. Grade I and II rhymes), we do not have enough understanding to account for more complicated phenomena such as the development of *ɣ- before two medials (i.e. Grade III and IV rhymes).

A rhyme is written with a hyphen before the vowel, e.g. *-a(C); a coda is written with a hyphen and a consonant, e.g. *-k.

3.2.5 Tone categories

There are four tone categories in Middle Chinese: *Ping*, *Shang*, *Qu* and *Ru* (You 2016: 92). These categories do not give us the actual tone contour. Moreover, as a reminder, the *Ru* category tells us whether the rhyme group consists of a stop coda.

Other important concepts applied in Chinese dialectology include the *Yin* and *Yang* registers. If the present-day dialect has experienced MC obstruent devoicing, the voicing of the MC initials (including sonorants) would determine which register the word fall into (Zhu 2016: 356, You 2016: 87). Assuming all tone categories split into *Yin* and *Yang*, the word which used to have a MC voiced initial would fall into the *Yang* register, otherwise *Yin* (You 2016: 87). Some dialects may have collapsed the distinction between the tone categories and/or the *Yin*/*Yang* registers due to mergers. A number of Yue dialects tend to preserve these distinctions[10] better than, for example, Beijing Mandarin, which is known for the loss of the codas in the *Ru* category (Yan 2006: 83).

There are correspondences between the MC tonal categories and the tone contours in present-

[10] See (2) in Section 4.1

day dialects. For example, *古* 'ancient' and *好* 'good' are in the *Shang* category in Middle Chinese. In Beijing Mandarin, these two words are pronounced with the tonal contour [^{214}]. This means the tonal reflex in the Beijing dialect for *Shang* is [^{214}]. Note that not all the words in the present-day dialects fully correspond to the *Guangyun* tonal categories. For instance, in Guangzhou Yue, the *Yang Shang* tone reflex is [^{13}] and the reflex for *Yang Qu* is [^{22}]. The word *道* 'road' belongs to the MC (*Yang*) *Shang* category, but it is pronounced as [tou^{22}] (i.e. it reflects the *Yang Qu*). This particular sound correspondence is called *濁上歸去* (*Zuo shang guei qu*, literally means voiced *Shang* belongs to *Qu*) in Chinese historical phonology. This is an example of a word that differs from the MC tonal category.

4. The Yue Dialect

Yue, together with other 'Chinese dialects', are groups of Sinitic varieties (Li & Thompson 2011: 703) in the Sino-Tibetan language family (DeLancy 2011: 693) or the Trans-Himalayan family (van Driem 2014). Yue is mainly spoken in the Guangdong and Guangxi provinces, as well as in Hong Kong and Macau. Yue is also spoken by many Chinese diasporas outside China, e.g. in Malaysia, Canada, the U.S. (Wu 2012: 125). It has been estimated that there are around 6,800,000 Yue speakers in the world (Wu 2012: 125).

In this section, the classification and internal classification of Yue are briefly overviewed, followed by three theories of the formation of Yue and the use of 'dialect' in this dissertation.

4.1 The classification of Yue

During the 20th century, many Chinese dialectologists have proposed different classifications of Sinitic varieties based on the reflexes of Middle Chinese features. Wang's (1936) *Chinese phonology* divides Chinese dialects into 5 main groups. Yue are the dialects which a) do not have voiced initials, b) possess codas -m, -p, -t, -k, and c) have at least 7 tone categories. It has been mentioned in Section 3.2.2 that Middle Chinese has voiced initials. As a result of MC obstruent devoicing, Yue dialects do not retain any MC voiced initials. In the last paragraph of Section 3.2.4, I mentioned that MC syllables could have stop and nasal codas. These codas are listed below in (2). In Yue dialects, these MC codas are all retained, according to Wang. In terms of tone categories, there are more than 4 categories due to the split into *Yin* and *Yang* registers (see Section 3.2.5). It is important to note that Wang's classification was based on limited Yue dialect data. The later discovery of more dialect data shows that not all Yue dialects fit the description above.

(2)

Stop codas	-p	-t	-k
Nasal codas	-m	-n	-k

Yuan (2001) divides Chinese dialects into 7 groups, but Yuan does not state how exactly he classified them. Zhou & You (2019: 9) also divide Chinese dialects into 7 groups, as shown in their tree diagram, given here as Figure 4. The diagram shows when each group split from the Northern Dialect in chronological order, as well as the closeness between dialect groups. In Figure 4, Yue is shown to be a separate dialect group that split from Northern Chinese after Wu and Xiang (Zhou & You 2019: 9). An alternative proposal comes from the *Language Atlas of China* (*2nd Edition*) (CASS & CityU LISRC[11] 2012), which Yue belongs to one of the 10 major dialect groups of Chinese.

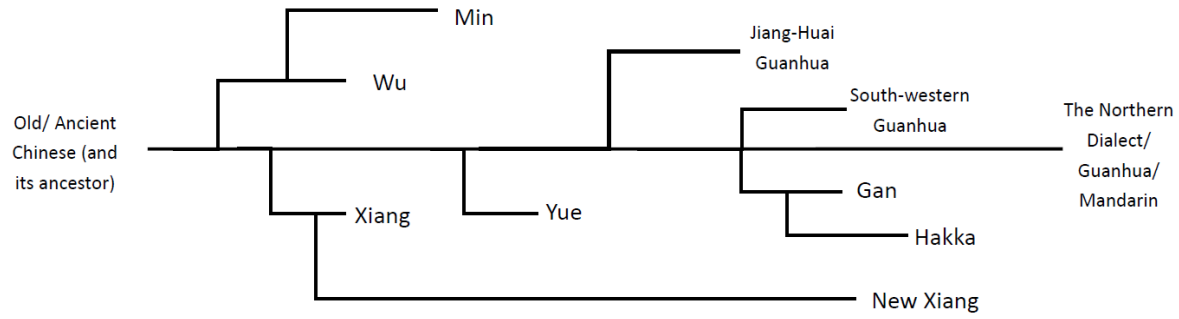

Figure 4. Tree diagram of Chinese dialects (translated from Zhou & You 2019: 9)

4.2 The internal classification of Yue

Yue is not a single homogenous variety that is spoken in such a wide area, unsurprisingly. In the last 60 years, there have been several attempts by dialectologists to classify the dialects of Yue. These attempts are either based on a selection of features chosen by the dialectologist or based on extralinguistic factors (e.g. geographical regions). Some of the classifications also do not cover the whole Yue-speaking area.

Zhan (1981) divides Yue into 5 different groups in his introduction of Yue varieties. However, Yue-Hashimoto (1991) argues that this classification is just a more detailed description of the

[11] Chinese Academy of Social Science and City University of Hong Kong Language Information Sciences Research Centre.

classification that Yuan (1960) proposed, hence not a novel classification. In Yuan's proposal, dialect groups are based on regions (different areas within the Guangdong and Guangxi province), without any linguistic justification. Later, Zhan (2002) describes features of 8 sub-dialect groups of Yue in Guangdong in detail and also compares the dialects within the group and with the Guangzhou dialect (the representative dialect of Yue). The author did not state explicitly the criteria he used to find each groups.

Xiong's (1987) classification can be found in the linguistic map of Guangdong in Figure 2 (see Section 2.3). He proposes a classification of Yue in Guangdong based on 3 features: a) the developments of Middle Chinese voiced stops; b) the reflex of MC *t^h is an [h] and c) the reflex of MC *s- is a [ɬ-]. By examining these features in this specific order, the result can yield 5 dialect groups for Yue. Based on a), obstruents dialects are divided into 3 groups: the Goulou dialects in which dialects possess unaspirated voiceless reflexes of MC voiced obstruents (across all tone categories); the Wu-hua dialects, in which all reflexes of MC voiced obstruents are voiceless and aspirated (across all tone categories), and the remaining dialects form a third group. Within remaining dialects, varieties that have feature b) are grouped as the Siyi dialects. The yet unclassified dialects are then classified into two groups: Guangfu dialects are dialects without feature c) and Gao-Yang dialects are the dialects with feature c). What I have just described is illustrated in Figure 5.

A major problem with Xiong's (1987) classification is that they are based on features that were cherry-picked in order to get specific results. In other words, choosing a different set of variables or switching the order in which groups of dialects are eliminated first might yield a completely different classification. For example, if feature c) is the first feature to be analysed, then some Siyi dialects (Taishan, Kaiping and Heshan) and the Gaoyang dialects would be grouped together, but the rest of the Siyi dialects, namely Doumen, Xinhui, Jiangmen and

Enping, would be left out. If I choose a completely different feature as a criterion, e.g. i-diphthongization in earlier Yue (Li 1997: 426-427), the classification would also be significantly different too. What is now the Guangfu dialect group would be split into fragments because i-diphthongisation is not found with geographical continuity.

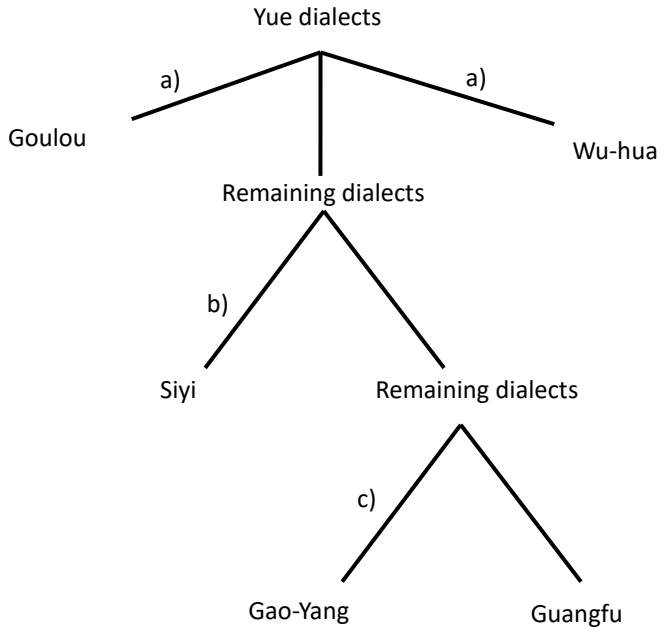

Figure 5. Classification of Yue dialects in Guangdong (Xiong 1987)

Lastly, Yue-Hashimoto's (1991: 167) classification is based on 167 features on tonal, initial, rhyme and lexical features and 63 dialects were investigated. It is the analysis which examined the most features so far. Unfortunately, due to the lack of data, this classification does not cover some Yue-speaking areas in Guangdong, for instance, north-western Guangdong (Yunan, Yunfu, Luoding, Fengkai, Guangning etc.). Hence her classification will not be referred to in this dissertation.

4.3 The formation of Yue dialects

There are three main theories on when and how Yue dialects were formed. A brief overview for each theory is given in the subsections below.

4.3.1 Yue was formed during the Qin dynasty

Yuan (2001: 178-179) and Zhan (2002: 3) have a similar view in the formation of Yue. They believe that the Qin army of 500,000 men brought to the south to settle a rebellion[12] caused an increase of the proportion of Han people compared to the local indigenous population. These men also brought their culture and their own dialect varieties to the area by settling down after they put down the rebellion. At the same time, Baiyue became part of the Chinese Empire, which means that it was under Chinese rule and the empire's influence. These factors caused the indigenous population to become more and more assimilated with the Han people. Another wave of migration happened during the Three Kingdom period. People fled from the north to the south due to war. A portion of the migrants fled to present-day Guangzhou. The growth in the population and economy increased the contact between the Han people and the indigenous people. This consolidated the status of Chinese in the area, which was significant for the development of Yue. Zhan calls this contact variety between northern Chinese and the indigenous languages the Early Yue dialect, which got increasingly more assimilated to northern Chinese and yet preserved features imposed from the ancient Yue indigenous languages. Zhan believes that the Yue dialect emerged as a unique dialect within the Sinitic family from this point onwards. Yuan believes that the mountains surrounding Guangdong acted as a communication barrier between the north and the south. This led to the divergence of this transplanted northern variety.

We do not have any documentation of the indigenous languages (which are believed to be Tai-Kadai languages, Yuan 2001: 179) spoken in the area before the arrival of the Han people.

[12] See Section 2.2.

However, through a comparison to present-day Zhuang, a Tai-Kadai language spoken predominantly in Guangxi, some similar linguistic properties can be found between present-day Yue and Zhuang. For example, the Siyi dialects (e.g. Taishan, Xinhui, Enping and Kaiping dialects) have the segment [ɬ], which is also found in Zhuang (Yuan 2001: 179). Some Yue-specific words such as [nɐm] 'think', [liu] 'play', [na:m] 'belly' are not found in other Chinese dialects, but can be found in some Zhuang dialects (Yuan 2001: 179). These features suggest contact between the two varieties.

Both Yuan's (2001) and Zhan's (2002) hypotheses point to the same direction that a) Yue dialects emerged as a result of migration from the north during the Qin dynasty and b) Yue was formed from language contact between the northern Chinese varieties at the time (Old Chinese) and the local indigenous languages. There are demographic and some linguistic evidence for this hypothesis. One should be cautious with the linguistic evidence provided, though. Even if Yue shares some phonological and lexical similarities with Zhuang, this could be evidence of contact that occurred in a later period instead of the Qin dynasty.

4.3.2 Yue was formed at the end of the Tang dynasty

Mai's (2009) account begins with the sociolinguistic situation in the Lingnan area in ancient times. He suggested the earliest form of Yue could have formed from language shift and early dialect contact between northern Chinese (from the migrants) and the indigenous interlanguage.

However the earliest form of Yue seems to have little to do with present-day Yue. Mai argues that present-day Yue is much closer to Middle Chinese than Old Chinese (Chinese spoken during the Qin dynasty). Therefore, one could rule out that Ancient Lingnan Chinese survived until today with endogenous changes in isolation. Instead, Yue is likely to be the product of the bombardment of northern Chinese influence for centuries. Mai introduces the concept of linguistic *strata*.

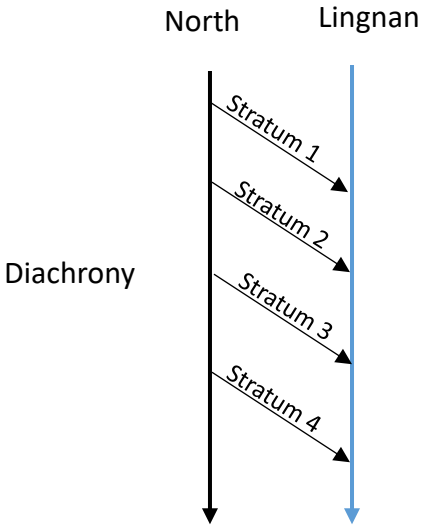

Figure 6. Influence from the Koine to Lingnan Chinese across time (modified from Pan 2004)

Lingnan is located at the peripheral area of the cultural centre, which was in the north. Therefore, Lingnan had been consistently under influence from the north, including linguistically. This is visualized in Figure 6. 'North' represents the cultural centre. The Chinese variety spoken there is called the 'Koine'[13] by Mai. The vertical arrows indicate the diachronic development of two varieties, one spoken in the North (the Koine) and one in Lingnan. The slanted arrows represent waves of influence from the Koine to Lingnan Chinese. These waves are also known as *strata* in the literature. Figure 6 illustrates that Lingnan Chinese had its own developments as well as receiving influence from the northern Koine. Lingnan Chinese is by no means a copy of the northern Koine in different periods (Mai 2009: 223). It is important to note that the newer stratum swamped the older stratum/-a. The phonological system of present-day Yue is the result of several layers stacked together.

Following Mai's theory of stratum-swamping, this predicts that whichever stage of the koine matches the structure of present-day Yue reveals when the latest stratum came. Mai discovered that the late Tang stratum was the latest stratum that completely swamped Yue. In other words,

[13] 'Koine' here refers to the 'common language'.

the basis of Yue was formed in the late Tang dynasty.

Mai's proposal depends on the linguistic evidence from present-day Yue. The transliteration of Sanskrit in the Northern Song dynasty (the dynasty after Tang) suggests that *ʊ was the reflex of the MC *Wei* initial[14] (Mai 2009: 224). Since Yue does not have this innovation at all, the stratum of the Koine that Yue corresponds to should predate the Northern Song dynasty. Another piece of evidence comes from the book *Huang ji jing shi shu* (Mai 2009: 224), written by Shao Yong (1011-1077 CE) during the Song dynasty. Some sections in this book are descriptions of the phonological system of the variety at the time (Zhu 2016: 421). It shows that syllables with historical *-t and *-k already merged with open syllable categories. Again, this innovation is not found in Yue[15]. Another evidence points to Yue's formation towards late Tang, but not before. Again in *Huang ji jing shi shu*, the MC palatal/retroflex stop series[16] shows signs that affrication has already taken place in a small number of lexical items (Mai 2009: 224). In Yue, the reflexes of these palatal stops are all affricates. What Mai seems to imply here is that the koine that swamped Yue seemed to have split off around the time this book was published. If the stratum that Yue corresponds to split off before this period, affrication would not have initiated[17]. Therefore, Mai (2009: 224) believes that the stratum Yue corresponds to cannot predate late Tang. Mai concludes that these features can show that Yue corresponds to the stratum of the koine spoken between late Tang and Northern Song dynasties.

In terms of historical evidence for Mai's theory, he points out that at the end of the Tang dynasty, the Lingnan area was relatively stable (Mai 2009: 224). Between the Tang and Song dynasty, there was a period called the Five Dynasties and Ten Kingdoms period. During this time,

[14] It is commonly reconstructed as *ŋ in Middle Chinese.
[15] In Section 4.1, it has been mentioned already that Yue has been conservative in terms of the coda inventory.
[16] See Section 3.2.2.
[17] Mai's theory seems to have rejected the possibility that affrication of MC palatal stops could be an independent development of Yue.

present-day Guangdong and Guangxi belonged to one of the kingdoms, known as the Southern Han Kingdom. During this 70-year period, people in Lingnan were relatively isolated from the turbulent north. Mai thinks Early Yue was formed during this period (Mai 1993, 1997, cited in Mai 2009: 224). Early Yue started out being similar to the northern Koine, but it started to diverge and became a 'dialect'[18] as time went on (Mai 2009: 224-225).

Although Mai proposed that the late Tang stratum was the main contribution in the formation of Yue dialects, he also pointed out the stratum that came later in the Southern Song dynasty was also influential, though not enough to cause swamping (Mai 2009: 225-226). This is supported by linguistic as well as genealogical evidence. For instance, the reflex of the Grade III vowel in the MC *Jia* rhyme group is [ɛ] and the merger of literary pronunciation of the MC *Zeng* and *Geng* rhyme groups are innovations in the koines spoken from the Southern Song dynasty onwards. These two features are also found in Yue. Based on the linguistic evidence alone, it is unclear when this stratum of influence reached Yue. Mai provides genealogical evidence which points towards Southern Song. The genealogy books and oral family history show that a massive number of families originally came from northern Guangdong, in Nanxiong Zhujixiang. The time that these families moved out from Zhujixiang was around the end of the Southern Song dynasty.

The last remaining question is why the Southern Song stratum did not swamp the late Tang stratum. Mai's argument implies that Yue was already independent as a 'dialect' after the Five Dynasties and Ten Kingdom period. The influence from the koine at that time has already reduced. Hence, the migration during the Southern Song dynasty then only caused some linguistic changes, not swamping of the Tang stratum.

[18] Topolect is also a more appropriate term here.

4.3.3 Yue was formed at the end of the Southern Song dynasty

Looking at the migration history of Guangdong, You (2016: 138) believes that the migration from northern China to Guangdong at the end of the Song dynasty (960-1279 AD) essentially formed the basis of present-day Yue. The correspondences between Yue and Middle Chinese seem to match the migration history. Lau (2001: 188-200) looked at the migration history to Guangdong in more details and at the same time found linguistic evidence which supports the same conclusion.

Unlike some other scholars, Lau (2001) focused on the formation of Yue varieties that are spoken around the Pearl River Delta, Xijiang and Western Guangdong. Despite dialectologists calling a number of varieties 'Yue', their actual linguistic structures are quite heterogenous. Lau is fully aware of this situation and therefore his generalisations do not include the more heterogenous dialects outside the area he focuses on.

	Tang			Northern Song			Southern Song	Ming
CE	639	742	820	980	1080	1180	1230/ 1252	1393
Guangdong	-	285456	149139	92559	579253	513171	445906	386430
Guangzhou	12463	42235	74099	16687	143261[19]	185713	170216	165220
Shaozhou	6960	20764	9666	10756	57438	-	19584	7810

Table 6. Population of Guangdong after the Tang dynasty (Liang 1980: 139, Lau 2001: 191)

Tables 6 shows the population data from the historical records (Liang 1980, cited in Lau 2001: 191). Towards the end of the Tang dynasty (820 CE), the population in Guangdong had dropped massively (compare to the population in 742 CE). However, at the beginning of the Northern Song dynasty, the population of Guangdong had risen almost six times from the original

[19] 55% of the population were immigrants. Therefore, around 65,000 people originally settled there

population (from 980 to 1080 CE). Although the population dropped slightly in the next 100 years, the overall population was still quite stable (compare to the transition between Tang and Song, Lau 2001: 191). Wu (1997) points out that the majority of the immigrants in Guangdong during the Song dynasty originally came from the Jiangxi province (northeast of Guangdong) and the rest were from the provinces further north; these immigrants all spoke a variety of Chinese in the north. Lau believes that there was a dialect shift in the local communities in Guangdong because the northern Chinese variety had prestige over the local varieties. Therefore, present-day Yue (discussed in Lau 2001) did not directly descend from Early Yue, a variety which was formed during the Qin dynasty.

The linguistic features of Yue that Lau presented suggest Yue is a descendant of northern Chinese in the Song dynasty. Lau claimed that the initials in Yue show a lot of 'modern' features, which developed after the *Qieyun-Guangyun* systems[20]. For example, there is no voiced consonants in Yue; the MC *Xia* initial (*ɣ-) has disappeared before *-ua; the MC sequence *xu- became [f-], etc. In terms of the rhymes, there are still a lot of correspondences with *Guangyun*, but a lot of the contrasts had also collapsed in Yue. For instance, the rhyme groups *Liu* and *Tong* have no contrast with their Grade I and III nuclei. In terms of tones, they don't completely correspond to Middle Chinese anymore. This statement mainly applies to words in the MC *Yang Shang* tone category[21]. Around two-thirds of the words in this category have merged with the *Yang Qu* category. In terms of the lexicon in Yue, the majority of the words came from the Ming dynasty.

4.4 Yue as a 'dialect'

4.4.1 The Chinese definition of 'dialect'

Li (2007: 1) defines the term *dialect* as "a vernacular that is used in a particular region" (my

[20] See Section 3.2
[21] See the example given in Section 3.2.5

translation). This definition has been used traditionally in China and it still applies nowadays. "A particular region" can range from one single village to several provinces. 'Language' on the other hand, as an academic term, refers to "the language of the nation". This means, the national language, Standard Chinese, is above Mandarin varieties as well as all other regional varieties spoken in China. This is similar to the concept of a 'language' having its own 'varieties', which also includes regional 'dialects' (oppose to social 'dialects'). According to Li's definition, Yue is a (regional) 'dialect' of Chinese, a variety of Chinese.

4.4.2 Western definitions of 'dialect'

Chambers and Trudgill (1998: 4-5) recognise that the definition of *dialect* is dependent on what a *language* is, but the definition of *language* is often extralinguistic. Chambers and Trudgill suggest to use the term *variety* instead, which is more neutral compare to the term *dialect*. Chambers and Trudgill's discussion introduces a role for political factors (as well as all other factors mentioned). This seems to be the case for Yue's dialectal status.

Kloss (1967) coined the terms *Ausbau* and *Abstand* languages. An Abstand language is paraphrased as "language by distance" by Kloss (1967: 158). "Distance" refers to the "intrinsic" distance between the varieties in question (how distance is measure is not a concern here). An Ausbau language on the other hand is called "a language by virtue of its having been reshaped". What Kloss means here is that an Ausbau language has been involved in "deliberate language planning… so as to become vehicles of variegated literary expression" (Kloss 1967: 158-159).

Yue is an Abstand language when comparing to Mandarin/Standard Chinese. The intrinsic distance can be reflected in the mutual unintelligibility (Cheng 1997, Tang & van Heuven 2015) between these two varieties as well as objective distance measures (Cheng 1997). Yue in Mainland China is not an Ausbau language.

4.4.3 The usage of the term *dialect* in this dissertation

To conclude, Yue's recognition as a dialect is not a purely linguistic question. Most of the definitions of *dialect* in section 4.4 relies on *language* or the standard variety. In this dissertation I will not define whether Yue is a language or not. I will only address it as Yue. The vernaculars collectively known under Yue are addressed as dialects or varieties interchangeably. For traditional classifications as well as the dialect groups of Yue, I will follow the tradition and use the traditional terminologies.

5. The Middle Chinese sequence *ɣu-

5.1 Introduction

The Middle Chinese initial *Xia*, or *ɣ-, is often reflected as an [h-] in most varieties of Yue (Yan 2006: 208-209). At a pure surface level description of the change, this looks like it is a case of devoicing and debuccalisation. There could be some exceptions to this phenomenon in some dialects, e.g. in the Jiangmen (Baisha) dialect, *ɣ- is realised as [k-]. In other dialects, it might also be reflected as [k^h-] (Yan 2006: 208-209). Yan (2006: 209) has mentioned that the realisation of *ɣ- as [k-] and [k^h-] is a unique feature of the Min dialects, but it is also found in Gan, Hakka and Yue dialects.

The development of MC *ɣ- before the MC medial *-u- is different from what occurs in other environments. In Chinese dialectology, words with both the *ɣ- and *-u- are known as *匣母合口字* [words with MC *ɣu-].[22] I have given some examples in Table 7.

Locations	'to return'	'lake'
Middle Chinese	*ɣuan	*ɣo/*ɣu
Guangzhou	[wan]	[wu]
Taishan	[van]	[vu]
Dongguan	[van]	[fu]
Maoming	[ʋan]	[fu]

Table 7. Pronunciation of 'to return' and 'lake' in several Yue dialects and Middle Chinese

Table 7 illustrates some of the inter- and intra-dialect variation that exists in Yue. On one hand, the pronunciation of 'to return' shows that the reflex of the MC *ɣu- sequence ranges from [w-], [v-] to [ʋ-]. However, for the word 'lake', there is no instance of [ʋ-]; an [f-] is found instead. Furthermore, the reflex is not the same in every word for all the dialects. We can see in Table 7 that some dialects have [w-]/[v-] for both words, while other dialects have [v-]/[ʋ-] combined with [f-].

[22] See Section 3.2.3.

Given the variation we see above, the MC sequence *ɣu- invites detailed exploration. In Chinese dialectology, phonological correspondences between the dialect and Middle Chinese are often described, but no closer analyses are given. The correspondences between a reflex and a MC category are often stated without any reference to the MC sound value (e.g. in Zhan 2002). The sound changes involved in these correspondences are therefore open to interpretation. Thus, there is a very limited amount of studies which attempts to cover as many as 54 Yue dialects all at once. Historical phonological studies usually focus on major, well-known dialects, as described in Section 5.2.3 & 5.2.4 below. This leads to a very partial and fragmented understanding of the sound changes for *ɣu- and other areas of Yue dialect variation. To deepen the understanding of the phonological history of Yue dialects, one has to expand the scope to more dialects of a region and in greater detail.

Previous studies of MC *ɣu- are reviewed in Section 5.2 for a number of dialects and a detailed analysis of the Yue dialects in Guangdong can be found in Section 7.

5.2 Previous studies

5.2.1 General sound correspondences between Middle Chinese to Yue dialects

The *ɣu- sequence (found with *closed rhymes*[23]) shows complicated developments in different Yue dialects (see Table 7). Yan (2006: 210) formalises these sound correspondences[24] between Middle Chinese and Present-day Yue varieties as:

(3) *ɣ- → Ø-, w-/ #___ closed rhymes

(4) *ɣ- → f-, v-/ #___ closed rhymes

[23] See Section 3.2.4

[24] Yan (2006) calls (3)-(5) 'phonological rules' and 'developments', but she never explicitly states that these are the changes from Middle Chinese nor using the '>' arrow to indicate change.

According to Yan (2006: 210), (3) was a sound correspondence that is found in Guangzhou, Sanshui (Xinan) and Gaoming (Mingcheng) dialects, whereas (4) is found Siyi dialects. Yan provides some example for (3) and (4). 'household' and 'to protect' are pronounced [wu] (< MC *ɣo/*ɣu for both words) in the dialects in which (3) is found (Yan 2006: 209). (4) is found in the Siyi dialects; the word 'great' is pronounced as [faŋ] (< MC *ɣuəŋ) and 'China, splendid' is pronounced as [vua] (< MC *ɣua) in the Doumen (Doumenzhen) and Xinhui (Huicheng) dialects, other Siyi dialects pronounce it as [va]. Yan also listed data from the Guan-Bao dialects (e.g. Hong Kong Kam Tin and the Dongguan dialects) and they show that (4) is also present. Some lexical items do show a different change from (3) and (4). In numerous Siyi dialects, the word 'to row (a boat)' is pronounced as [p^ha] (< MC *ɣua). Yan (2006: 209-210) formalises this in the rule shown in (5) below.

(5) *ɣ- → p^h-/ #___ closed rhymes

5.2.2 An overview of the reflexes of MC *ɣu- in different Yue dialect groups

Zhan (2002) gives an overview for the characteristics of each dialect group that he proposed (see Figure 7) and one of the features that he describes is the reflexes of MC *ɣu-. This subsection is based on the description Zhan (2002) provides. The data that Zhan uses overlap with the data used in this dissertation, with some additional data that the author collected with his colleagues (Zhan 2002: 2). This description provides a good basis for the more detailed dialectal phonological analysis in this dissertation. I have summarised Zhan's descriptions in Figure 7 below.

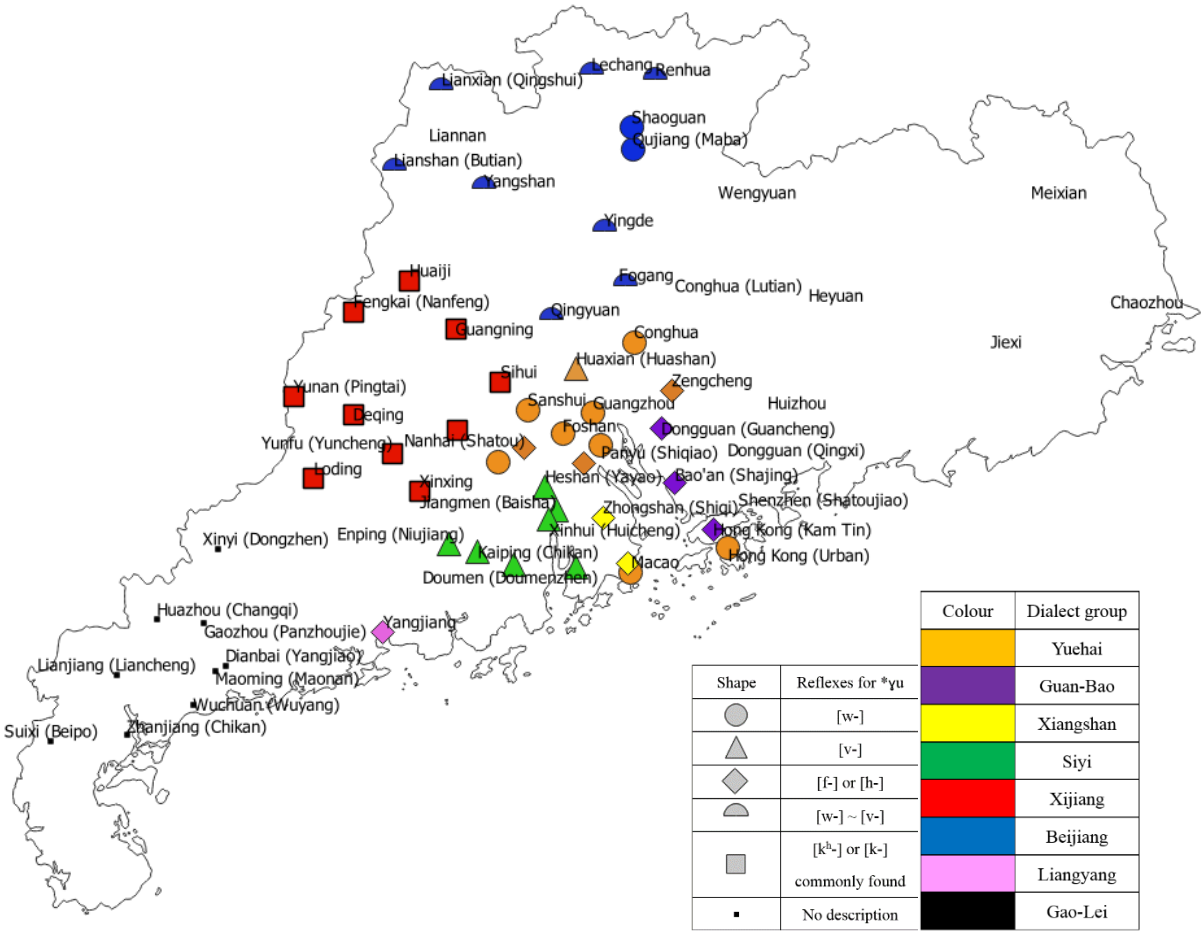

Figure 7. Sub-dialect groups of Yue (Zhan 2002) and their reflexes of *ɣu-[25]

The map in Figure 7 illustrates Zhan's (2002) general observation of the reflexes of MC *ɣu- in various Yue dialect groups. Colours are used in this map to indicate different dialect groups and symbols are used to represent the reflexes of *ɣu-. There is no description available in Zhan (2002) for the Gao-Lei dialect group unfortunately.

The Yuehai and Guan-Bao dialects around the Pearl River Delta generally reflect MC *ɣu- as [w-] (Zhan 2002: 127, 189), represented by circles in Figure 7. However, some reflexes are found to be [f-] in Guan-Bao and Xiangshan dialects (the Zhongshan dialect has [h-] (< *f-), Zhan 2002: 189, 199), as well as in three Yuehai dialects near Guangzhou (Zhan 2002: 127), indicated by the diamonds in Figure 7. Moreover, some dialects, namely Dongguan and

[25] This map and other self-made maps are produced with QGIS (QGIS Development Team 2020)

Huaxian, are found to reflect *ɣu- with [v-]. In the north, Zhan (2002: 163) observes that many dialects have the tendency to pronounce what we find as semi-vowels in the Guangzhou dialect as voiced fricatives; others have free variation between [w-] and [v-] (indicated by half circles in Figure 7) with the exception of Shaoguan and Qujiang dialects. Zhan does not provide any examples for this. In the Xijiang dialects in the west, Zhan (2002: 177) notes that [k-] and [k^h-] are commonly found as reflexes for MC *ɣ- and the *ɣu- sequence (indicated by squares). For example, the word 'fox' is pronounced as [wu] (< *ɣo/*ɣu) in Guangzhou, but in the Huaiji dialect, it is pronounced as [ku]. The Siyi dialects have been described as having [v-] as the reflex (see the triangles in Figure 7) and, lastly, Yangjiang has an [f-] as a reflex for some words.

5.2.3 Dialect-specific diachronic analyses: Guangzhou

Chen & Newman (1984a, 1984b, 1985) have written an extensive phonological history of the Guangzhou dialect (from Middle Chinese). It consists of three parts: initials, rhymes and tones. Their phonological history has an account of the developments of the MC sequence *ɣu-.

(6) $$*ɣ- \quad > \quad Ø \quad / \# \_\_ \begin{bmatrix} -cons \\ +hi \end{bmatrix}$$

Chen & Newman (1984a: 188) call (6) ɣ-procope. ɣ-procope essentially affects occurrences of MC *ɣ- which precede MC medials[26] *-u-, *-y- and *-i-. After this change applied, a glide could be inserted in these syllables (Chen & Newman 1984a: 188). There are two remarks that come with this change. The first is that ɣ-procope must "apply before DEVOIC[ing of *ɣ-]", i.e. ɣ-devoicing, the consequence of which would be the merger of MC *ɣ- and *x-. The other remark from Chen & Newman (1984a: 188) is that ɣ-procope applied before MC *o, but not MC *u.[27] This can all be seen in (7), taken from Chen & Newman (1984a: 188):

[26] See Section 3.2.3

[27] This is different from MC medial *-u-, which requires a nucleus following it. Here, *u is a nucleus vowel.

(7)

Character	候	互
Gloss	‘marquis’	‘mutual’
Middle Chinese form	*ɣu	*ɣo
Diphthongisation	ɣəu	-
Raising	-	ɣu
ɣ-procope	-	u
Glide-insertion	-	wu
Devoicing	həu	-
Present-day form	[hɐu]	[wu]

(7) illustrates the relative chronology of a few sound changes. It shows that the rhyme *-o did not trigger ɣ-procope in Middle Chinese like a medial *-u- until the vowel had been raised. Chen & Newman’s analysis suggests that, when ɣ-procope was active, the result of diphthongisation of MC *-u has bled ɣ-procope while the consequence of MC *o-raising has fed ɣ-procope, according to Chen’s (1976a) reconstruction of Middle Chinese. To account for Chen & Newman’s analysis, I have shown their reconstruction together with Zhu’s reconstruction for the *Yu* rhyme group, which ‘mutual’, ‘lake’ (example in Table 7) and several other words belong to (see the *Yu* rhyme group in Table 9 in Section 6.3), i.e. *ɣo (Chen & Newman 1984a) and *ɣu (Zhu 2016). It is necessary to show these two stages of this class of words because, as I will show in Section 8.2, it can explain the historical changes which are responsible to the present-day dialect landscape.

It should be noted that words with the sequence *ɣu- should instead be written as *$ɣ^{w}$- (Chen & Newman 1984a: 188). The authors added that “the disappearance of the *ɣ- [as procope] amounts to a simplification of the initial *$ɣ^{w}$- rather than deletion of the initial”. This quote suggests that the change from *ɣu- to [w-] could be somewhat more gradient than what we see in (7). For instance, like this: *ɣu > *$ɣ^{w}$ > *^{ɣ}w > w. While this could be the case, the analysis remains categorical as the nature of the dialect survey data does not allow a phonetic analysis.

Another diachronic analysis comes from Chen (2011), in which the author uses dialectal evidence from the wider Guangzhou area as well as some Guan-Bao dialects for his diachronic analysis.

Chen (2011: 96) lists all the closed rhymes that have a zero reflex [Ø-] for MC *ɣu-, see (8). These words happen to match Chen & Newman's (1984a) ɣ-procope description.

(8)

Middle Chinese	Guangzhou example	Gloss
*ɣuɑ	[uɔ]	grain
*ɣua	[ua]	Chinese
*ɣo/*ɣu	[u]	lake
*ɣuɑi	[ui]	to return
*ɣuɑn	[un]	to change/ switch
*ɣuən	[uɐn]	spirit
*ɣuɑŋ	[uɔŋ]	yellow
*ɣuɐŋ	[uaŋ]	horizontal

In general, Chen (2011: 96) states that there are only two words, listed in (9), in which the reflex [f-] is found:

(9)

Middle Chinese	Guangzhou example	Gloss
*ɣuɑn	[fun]	pillar
*ɣuɑŋ	[fɔŋ]	to sway

Chen (2011: 96) thinks that the lexical items in (9) are infrequent words in the spoken language, therefore, they should not be included in the generalisation that the Guangzhou dialect has a zero reflex [Ø-] for MC *ɣu-, i.e. words reflecting MC *ɣu- with an initial [u-]. Chen & Newman's (1984a) ɣ-procope seems to be completely regular in Chen's (2011) Guangzhou

data, excluding ‘pillar’ and ‘to sway’.

However, Chen (2011: 97-98) finds that the description above does not apply to rural dialects spoken in the wider Guangzhou area. This includes the Shawan, Dagang, Xintang and Licheng dialects. In these dialects, the reflex [f-] is found in the pre-u# context. [f-] is also found in other environments, e.g. perceding [-ɔ], [-ui], [-un], [-ɔŋ] and [-ɔk]. The words with these environments overlap with the words listed in (8). Chen (2011: 99) states that the origins of the [fu] pronunciation comes from (*ɦu-[28]>) *hu > fu, which is a sound change that is found across Yue dialects, such as the Dongguan and Hong Kong (Kam Tin) dialects (Chen 2011: 101).

Regarding the chronological relationship between the zero reflex [Ø-] and the reflex [f-], Chen (2011: 102) cites two opposing theories in the literature to explain the appearance of [f-]. Lau (2003) argues that in Yue, the zero initial [Ø-] is found more often than other initials. He believes that this zero initial is a retention from “the pronunciation before the Tang dynasty”. This theory comes from an idea that Wan (1995) proposed (Lau 2003). In the Anyi dialect (a variety of Gan), there is a zero reflex for MC *ɣ- found before MC closed rhymes (Grade I, II), like Yue. Wan’s (1995) theory states that the Anyi dialect lost *ɣ- before closed rhymes in *colloquial pronunciation*, which reflects the endogenous changes of the dialect (cf. *literary pronunciation*, which reflects a borrowed pronunciation, Li 2007: 93). This theory predicts that [f-] is the *literary pronunciation* imported from Guanhua. The author states that “phonetically speaking, it is very implausible that the zero initial changed into an [f-]” (Wan 1995: 233). Borrowing therefore is a legitimate explanation for the presence of a reflex [f-] in Wan’s theory. Moreover, the zero reflex is also found in other more conservative Sinitic varieties in the south

[28] The reconstruction of Middle Chinese employed by Chen (2011) is different from Chen & Newman (1984a). Chen (2011) uses the reconstruction in Pulleyblank (1984), Pan (2000), which the MC initial *Xia* is *ɦ-. Also see Section 3.2.2.

(Min, Hakka). Wan (1995) argues that the conservative nature of dialects sharing the same feature provides support for the theory that the zero reflex is a reflection of the earlier stage of the language.

Lau (2003) applies Wan's analysis of the Anyi (Gan) dialect to Yue. This means that the reflex [f-] found in Yue is also seen as borrowings from Guanhua for Lau, and it was a development much later than the loss of *ɣ- (Lau 2003).

The opposing theory comes from Li (1996), cited in Chen (2011: 102). Li (1996: 75) claims that MC *ɣ- preceding a closed rhyme was lost and that [w-] became the new initial. For the words with an [f-] initial, e.g. 'to sway' [fɔŋ] (< *ɣuɑŋ), Li (1996: 76) believes that *ɣu-/*ɦu- firstly became *hu- through devoicing (and debuccalisation), then turned to [f-]. Li's theory does not suggest that words with [f-] are borrowings, but perhaps exceptions to the general sound change that he described earlier in the paper.

5.2.4 Dialect-specific diachronic analyses: Dongguan

Li (2015: 84-91) explores the labio-dentalisation of MC *ɦu-[29] in the dialects spoken in the Dongguan region. One of the most distinctive differences between the Dongguan dialects and the urban Guangzhou dialect is the present-day pronunciation of MC *ɦu (Li 2015: 87). Some examples (from Li 2015: 87) are given in (10) below:

(10)

	Guangzhou (urban)	Guancheng	Houjie	Shilong
'lake'	[wu]	[fu]	[fu]	[fu]
'fox'	[wu]	[fu]	[fu]	[fu]
'mutual'	[wu]	[fu]	[fu]	[fu]

[29] Li (2015) also employs [ɦ-] as the reconstruction of the Xia initial.

(10) shows that all the Dongguan dialects have the [fu] pronunciation for those three words whereas the Guangzhou dialect does not have the [f-] initial. Li (2015: 87) states that this pronunciation for the pre-u# words is also very common in non-Guangzhou Pearl River Delta Yue dialects as well as the Siyi dialects. Li (2015: 87) states that this group of words have experienced labio-dentalisation *h- > *ɸ- > f- (after ɦ-devoicing).

Although [f-] is found as a reflex in the pre-u# context, in other environments a zero reflex [Ø-] for MC *ɦu- is found, i.e. before a MC medial *-u- (Li 2015: 88). Li describes this as "the same as the Guangzhou dialect" (Li 2015: 88). Li (2015: 88) claims that the zero reflex occurs in words that are more frequently used in the spoken language; he suggests that a frequency effect has caused ɦ-devoicing in *ɦu-, then the loss of *h- (< *ɦ-). According to Li (2015: 88), this change occurred relatively long ago.

It should be mentioned that ɦ-devoicing leads to a merger of the reflexes of MC *hu- and *ɦu- to *hu-. Li does not explain how h-loss only affected the reflex of *ɦ- but not for *h- after the merger. His explanation is therefore not adequate to explain the variation of the reflexes of *ɦu- in the Dongguan dialects.

5.2.5 Implications of previous studies for this dissertation

Both Zhan (2002) and Yan (2006) provide an overview of the geographical patterns of the reflexes of MC *ɣu- found in Yue dialects. Chen & Newman's (1984a) study shows that MC *ɣu- went through ɣ-procope before a medial in Guangzhou Yue. The authors made a distinction between the MC medial *-u-, MC vowels *-o# and *-u#, which is very important in their analysis, because (early) MC *-u which went through diphthongisation of this *-u which bled the environment for ɣ-procope. The raising of MC *-o (> [u]), however, fed ɣ-procope. Chen's (2011) study on the other hand shows that some rural Guangzhou dialects

possess the reflex [f-] which could be a result from borrowing from Guanhua (Lau 2003) or having exceptions to the change (Li 1995).

Furthermore, both Chen (2011) and Li's (2015) analyses enhance our understanding of the change from ɦ-devoicing of *ɦu to labio-dentalisation of the *hu (< *ɦu) sequence. They argue that the [fu] pronunciation comes from *ɦu > *hu > *ɸu > fu. This reconstruction can aid the analysis of the changes in Yue dialects other than those of rural Guangzhou and Dongguan.

Based on these previous analyses and the descriptions, I have come up with a typology which captures the possible types of the reflexes and related changes that are present in the Yue dialects. The changes are shown in Figure 8 below. The processes in this typology are referred to in the analysis in Section 7.4 and and discussion in Section 8.

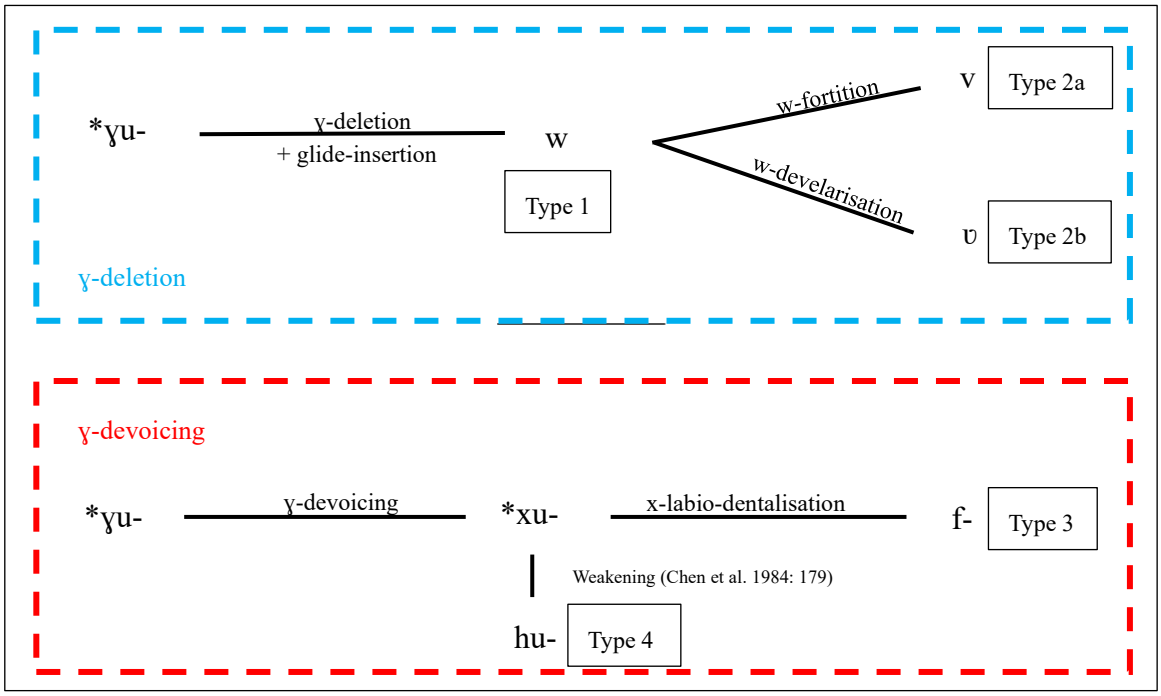

Figure 8. Typology of the development of *ɣu- and glide-insertion

Figure 8 gives a list of possible changes from Middle Chinese *ɣu- to present-day dialect reflexes in Yue. The top group of changes are related to *ɣ-deletion* (*ɣ-procope* in Chen &

Newman 1984a). Type 1 (ɣ-deletion and glide-insertion) is found in Guangfu dialects (Chen & Newman 1984a). Type 2 changes are secondary changes, because they are innovations which affect [w-] after ɣ-deletion. Type 2a (w-fortition) is found in Siyi dialects (Zhan 2002) and Type 2b (w-develarisation) is found in Zhan-Mao dialects (Shao 2016).

The bottom group of changes are related to *ɣ-devoicing*. Type 3 (x-labio-dentalisation) is found in Pearl River Delta dialects (excluding the urban Guangzhou dialect, Li 2015: 87), as well as the Yangjiang and Lianjiang dialects (Zhan 2002: 208, 211). Type 4 (weakening) is found in the Zhongshan dialect (Zhan 2002). Whether *ɣ-devoicing* is also found in other dialects is an issue that is addressed in the analysis in Section 7.4 and 8.

The changes listed in Figure 8 are assumed to be categorical. In this dissertation, no attempt is made to reconstruct and infer the history of phonetically-detailed changes. This is because, firstly, if these changes were really gradient, there could be an infinite amount of stages to be reconstructed. Listing every single instance of the changes involved does not make a meaningful contribution to the analysis. Secondly, the data was not collected based on a large population and no acoustic data was collected in the dialect surveys at all since no recordings were made. For these reasons, this dissertation only focuses on the categorical reconstruction of the development of the MC sequence *ɣu-.

6. Data

The data for this dissertation mainly come from four dialect surveys and one lexicon. The subsections below provide information for the informant selection process in Chinese dialectology as well as explaining the layout of these surveys and how data is extracted for this dissertation.

6.1 Informants

These surveys aimed to collect dialect data which represent the most local, externally-uninfluenced speech from each location, but some informants who took part in these surveys were slightly different from informants labelled as *NORMs* (Chambers and Trudgill 1998: 29) by western sociolinguists in a few ways.

6.1.1 Informant selection

The informants selected by traditional dialectologists are often described as *NORMs*. It stands for non-mobile, older, rural males (Chambers & Trudgill 1998: 29). There are good reasons for choosing NORMs as informants (Chambers & Trudgill 1998: 30):

Non-mobile: dialectologists may be interested in the speech that shows characteristics of the regions, and this kind of informant are more likely to provide examples of local speech.

Older: speech data provided by older speakers usually reflects more old-fashioned regionalisms.

Rural: dialectologists want to avoid as much mobility from the speakers as possible (less contact with other varieties); rural communities have relatively less mobility than urban communities.

Male: in the west, women's speech is thought to be more prone to class-consciousness and self-consciousness and women tend to lead change (cf. Gender Paradox, Labov 2001: Chapter 8). It has also been found (in England) that "men speak vernacular more frequently, more

consistently and more genuinely than women" (Orton 1962: 15). Therefore, male informants were preferred.

In contrast, the general guidelines for field workers in Chinese dialectology are listed below (Chinese Academy of Social Science (CASS) 1983: v, Li & Xiang 2010: 107, Li 2017: 16, You 2018: 17-18):

1) Informants have to be from the area (and stayed there for most of their lives),
2) Informants have to speak the basilectal, rather uninfluenced local vernacular,
3) Informants should have at least received secondary education for literacy,
4) Informants have to be middle-aged or older,
5) Informants should have no speech defects

We can agree that *non-mobile* and *older* are listed as criteria for finding an ideal informant in Chinese dialectology, but not for *rural* and *male* in NORMs. There is no mention of rural vs. urban dialects. In fact, the representative dialects of major Sinitic branches are all dialects from the biggest urban centres in their dialect area. Gender is also not restricted here. One can see in the Yue dialect surveys that numerous female informants were chosen. Educational background is also required for Chinese dialectology because the methodology for data collection is to ask informants to read out characters from the *Dialect Survey Wordlist*[30]. What can be certain is that Chinese fieldworkers also aimed to record the most local, uninfluenced dialects (what they described as basilectal local vernaculars), despite the differences in some of the criteria in informant selection.

6.1.2 Informants in the Yue dialect surveys

Since the surveys were conducted in the late 1980s and 90s (even in 2010s for one), it was harder to find older informants that were non-mobile for their whole life time, especially when so much has happened in the history of China in the 20th century (e.g. WWII, The Cultural

[30] *方言調查字表* in Chinese. See also Section 3.1.2

Revolution). Dialectologists however managed to find informants who did stay in the location where the dialect survey was conducted. Exceptions to this are usually for job-related reasons. In addition, many informants are also descendants of inhabitants of the area for many generations, e.g. a speaker of the Huaiji dialect belongs to the 23rd generation in the family living in Huaiji (Zhan & Cheung 1998: 811).

Age varies hugely for the informants in the dialect surveys. For example, in *Survey of Dialects in the Pearl River Delta* (*SDPRD*), the age of the informants for Yue dialects ranges from 41 to 80 (Zhan & Cheung 1987: 434-435). Zhan & Cheung (1987: 3) commented that these informants were capable of speaking "a relatively 'pure'[31] local vernacular". However, some informants were as young as 17 and 28 years old, e.g. in *Survey of Yue dialects in West Guangdong* and *Survey of Yue dialects in North Guangdong* respectively. Generational and age differences cannot be avoided while using these dialect survey data.

In the Yue dialect data, not all the localities were chosen from rural communities. An example is the Guangzhou dialect (from Zhan & Cheung 1987). Guangzhou is the biggest city in Guangdong. Another one is the Dongguan (Guancheng) dialect. In fact, there is a tendency that both urban and rural localities are chosen in various Yue dialect surveys. This is particularly obvious in the *SDPRD* since it contains the highest number of localities out of all the dialects surveys. The fieldworkers of the surveys (e.g. *SDPRD*) made judgements and chose informants that must be local and were capable of speaking the local vernacular (Zhan & Cheung 1987: 3). This might have reduced the influence from other dialects from mobility, which Chambers & Trudgill (1998: 30) mentioned.

[31] Presumably it means 'uninfluenced'

Last but not least, not all informants are male. As mentioned (see above in the same section), gender is not restricted in informant selection. It is unclear what kind of gendered speech variation existed in these dialects. It is also another aspect that cannot be avoided while using the survey data.

One other aspect of the speakers has not been mentioned yet. Some speakers were able to speak Guangzhou Cantonese, Mandarin and/or some other dialects. While these information has been noted in the informant's sociolinguistic background, we do not know to what extent the elicited data have been influenced by their competence of other Sinitic varieties.

6.2 Dialect surveys and other sources

This dissertation uses data from 4 dialect surveys and 1 lexicon.

6.2.1 *Survey of Dialects in the Pearl River Delta* (*SDPRD*)

SDPRD (Zhan & Cheung 1987, 1988, 1990) is a three-volume survey which was conducted between 1986 and 1987 (Zhan & Cheung 1990). The first volume consists of phonetic transcriptions of 3810 Chinese characters in 31 different dialects, the second volume is a collection of 1401 lexical items also in 31 dialects. The last volume is a summary of the data in an attempt to give an overall linguistic picture of the dialects surveyed in the project. The last volume consists of tokens of reflexes for each Middle Chinese sound category, dialect maps and speech samples from the informants. The phonetic data in this dissertation are taken from volume one.

In *SDPRD*, there are 25 Yue localities, 5 Hakka localities and 1 Min locality. Only the Yue localities are considered in this dissertation. The localities were chosen by counties or cities; for most of the Yue speaking counties, only one location has been selected for the survey. Figure 9 is a map of the survey points in *SDPRD* (Zhan & Cheung 1990: 321). The white circles

indicate Yue survey points; black circles indicate Hakka dialects and the stripy circle next to the Siyi area indicates the Min dialect spoken in Zhongshan. The map is in Chinese, but important locations such as Guangzhou, Hong Kong as well as important geographical areas, e.g. the Pearl River Delta, are indicated in English.

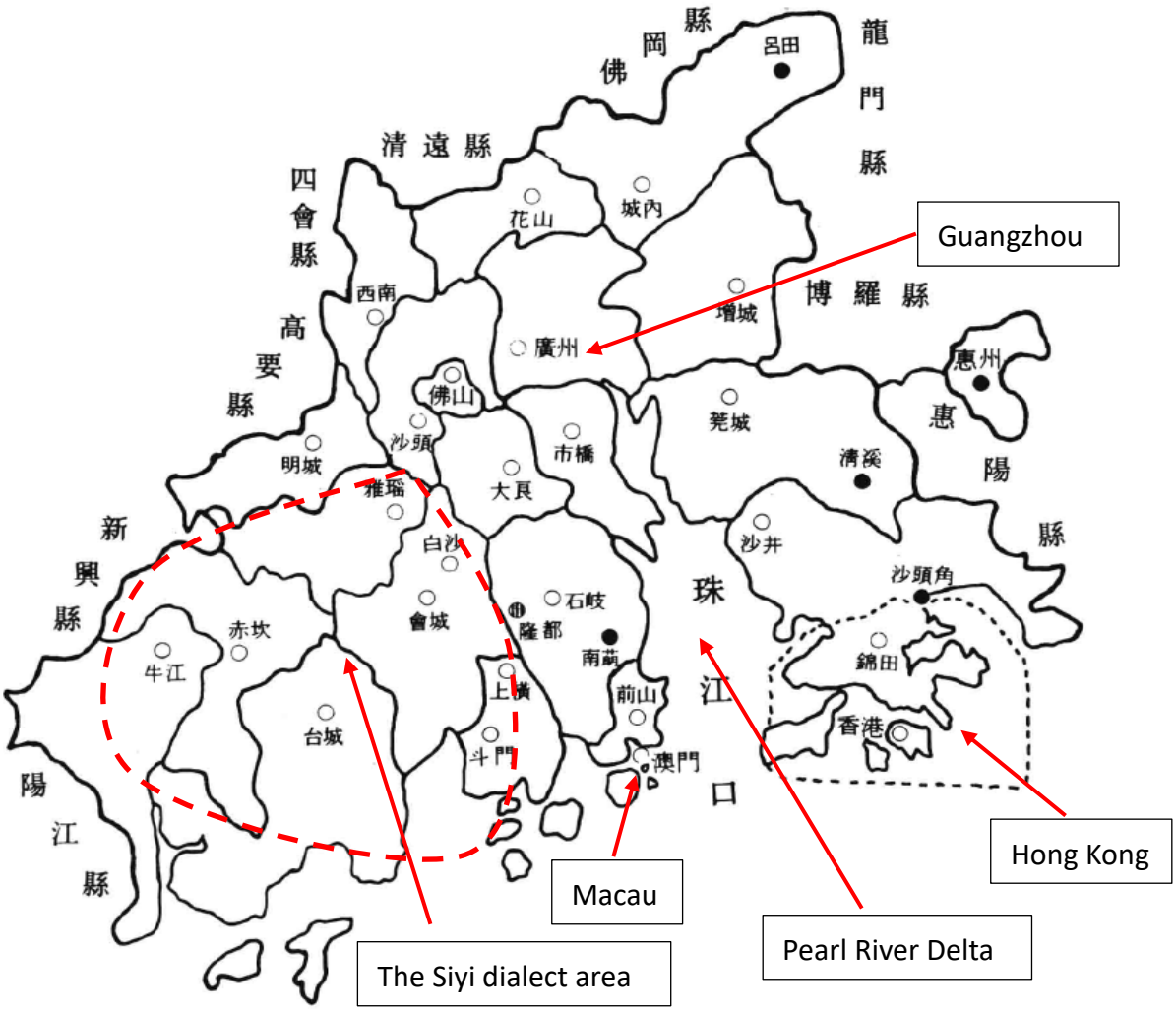

Figure 9. Survey points in *SDPRD* (Zhan & Cheung 1990: 321)

In this dissertation, data from the dialects listed in (11) are used:

(11) Guangzhou, Hong Kong (urban district), Hong Kong (Kam Tin), Macau, Panyu (Shiqiao), Huaxian (Huashan), Conghua, Zengcheng, Foshan, Nanhai (Shatao), Shunde (Daliang), Sanshui (Xinan), Gaoming, Zhongshan (Shiqi), Zhuhai (Qianshan), Jiangmen (Baisha), Xinhui, Taishan, Kaiping (Chikan), Enping (Niujiang), Heshan (Yayao), Dongguan (Guancheng), Doumen (Doumenzhen), Bao'an (Shajing).

6.2.2 *Survey of Yue Dialects in Northern Guangdong* (*SYDNG*)

SYDNG is part of a sequel to *SDPRD* (see Section 6.2.3 for the second part of the sequel), which is a Yue dialect survey conducted in northern Guangdong (Zhan & Cheung 1994: v). The aim for this survey was to expand the survey area of Yue in order to gain a fuller understanding of its dialects. This survey was conducted between 1992 and 1993.

The items in this survey are similar to *SDPRD*. There are phonetic transcriptions of 3810 Chinese characters. There is a difference to *SDPRD* which lies in the number of lexical items collected in the elicitation; only 1248 items were collected. This survey does not provide a detailed diachronic phonological comparison between the present-day dialects and Middle Chinese, unlike other dialect surveys (Zhan & Cheung 1994: viii). It purely serves as a database for future researchers to investigate dialects of this area. The first section of *SYDNG* contains the phonetic transcription of Chinese characters from the *Dialect Survey Wordlist*. This is the main source of the data for this dissertation for dialects spoken in northern Guangdong.

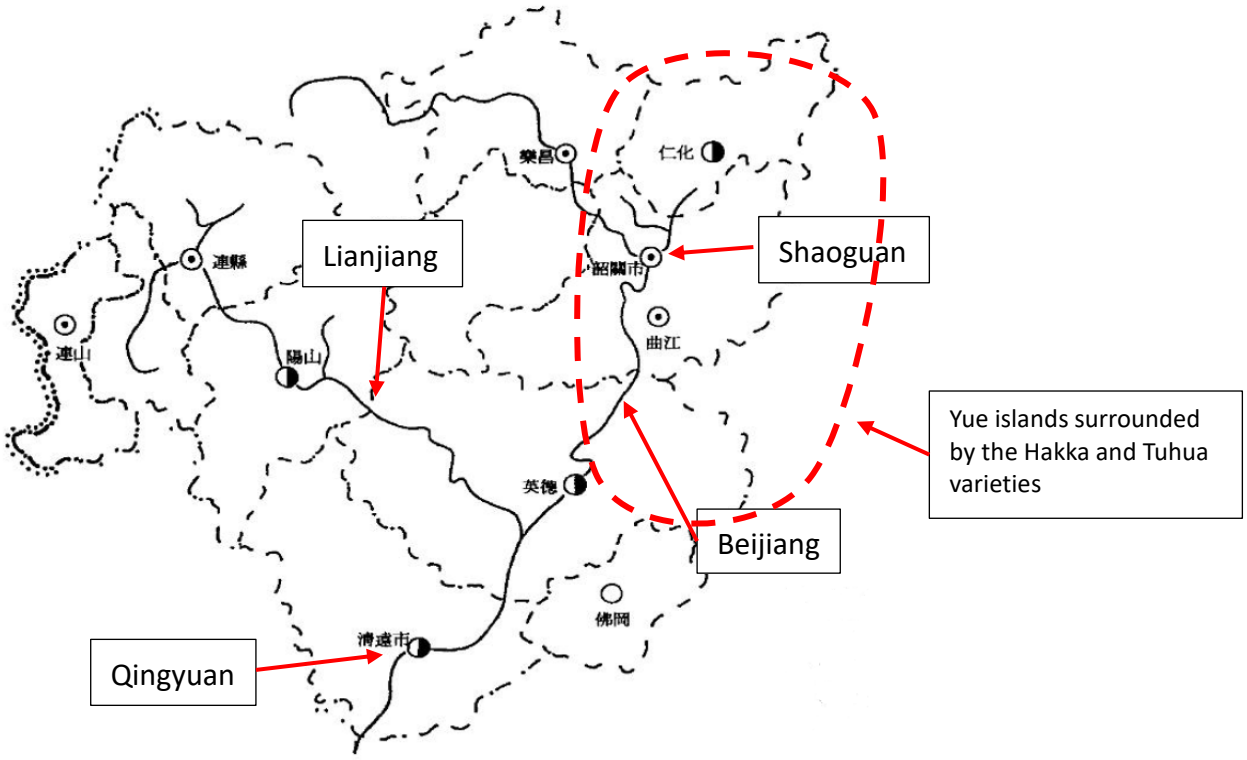

Figure 10. Survey points in *SYDNG* (Zhan & Cheung 1994: 688)

There are 10 survey points in total, all listed in (12), also see Figure 10 (from Zhan & Cheung 1994: 688). They are all located in northern Guangdong. The most important urban areas are indicated in English. The line that connects a few survey points together is the upper stream of Beijiang (North River). The northwestern stream in Figure 10 is known as Lianjiang (Lian River), which eventually merges with Beijiang. Lastly, the different symbols at each locality indicate the sociolinguistic situation. A white circle indicate a monolingual Yue-speaking community; a half black and white circle indicate locations that are bilingual in Yue and Hakka; a circle with a dot means the location is a multilingual community, where Yue, Hakka and some indigenous languages are spoken. All the localities within the red dotted circle in Figure 10 are Yue islands surrounded by Hakka and Tuhua varieties, on top of the descriptions above (see also the top of Figure 3 in Section 2.3). Lastly, this map in Figure 10 is located directly above the map in Figure 9.

(12) Qingyuan, Fogang, Yingde, Yangshan, Lianshan (Butian), Lianxian (Qingshui), Shaoguan, Qujiang (Maba), Renhua and Luochang.

6.2.3 *Survey of Yue Dialects in Western Guangdong* (*SYDWG*)

SYDWG is the second part of the sequel of *SDPRD* (Zhan & Cheung 1998: v). As a sister survey of *SYDNG*, the structure of *SYDWG* is pretty much the same as *SYDNG* (Zhan & Cheung 1994, see Section 6.2.2). This survey was also conducted between 1992 and 1993.

Like *SYDNG*, there are 3810 phonetic transcriptions of Chinese characters elicited from the informants in total. The comparative lexicon also only contains 1248 items. There are also transcriptions of spontaneous speech available for all dialects in this survey.

There are 10 survey points in total in *SYDWG*, listed in (13). They are all located in central-western Guangdong. Figure 11 below is a map of all the survey points in *SYDWG*. This map connects to the three counties in the central-western part of the map in Figure 9. Zhaoqing is a more widely known city of the area, therefore I have marked its location in English. Xijiang (West River) is also marked on Figure 11. In this area, Yue is the dominant Sinitic variety spoken, with about 80% of the population (Zhan & Cheung 1998: 3). There are also Hakka islands scattered across the counties in Figure 11, but a bigger cluster is found between Yunfu and Luoding (see the green circle). Min is also spoken in some isolated villages, but the speaker population was around 10,000 when the survey was conducted. It should be noted that minority languages such as Zhuang and Yao were used by small communities (around 4,000 people) near Huaiji.

(13) Zhaoqing (Gaoyao), Sihui, Guangning, Deqing, Huaiji, Fengkai (Nanfeng), Yunfu (Yuncheng), Xinxing, Luoding, Yunan (Pingtai)

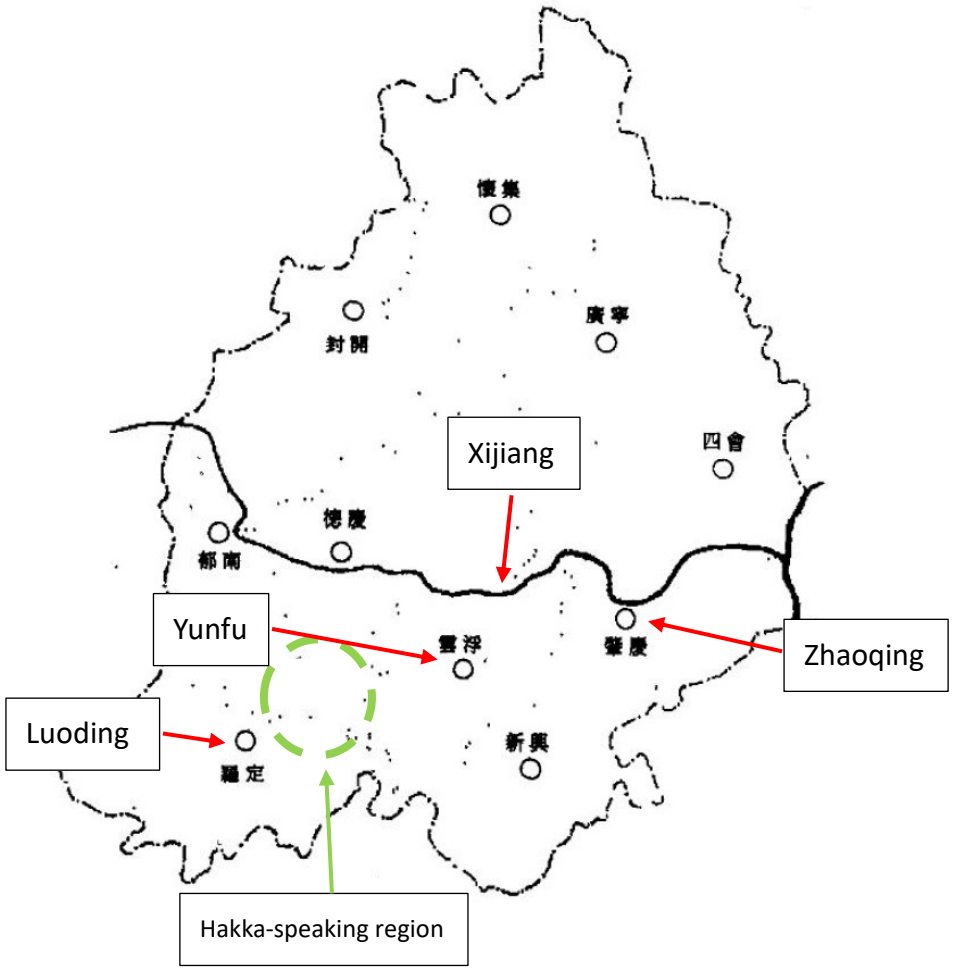

Figure 11. Survey points in *SYDWG* (Zhan & Cheung 1998: 693)

6.2.4 *The Phonological Study of the Yue Dialects spoken in the Zhan-Mao area in Western Guangdong* (*SYDZM*)

Despite the name, *SYDZM* (Shao 2016) is actually very similar to the dialect surveys introduced in previous subsections in nature. The aim of the book is to fill the gaps of dialect areas that were not covered by the earlier Yue dialect surveys from the 1980s and 90s as well as later studies (Shao 2016: 1-2). The survey was conducted between 2006 and 2011.

The format of this book is very similar to the other dialect surveys. The main content consists of the phonetic transcription of 3588 Chinese characters in 11 dialects in western Guangdong. In addition, there are 10 articles at the end of the book on phonological features of these dialects. The localities investigated in this dissertation are listed in (14) and are shown in Figure 12. The

region that this map shows is situated to the south-western part of the map in Figure 9, but not connecting to the south-western county on that map. These survey points are all located in (south-)western Guangdong, which does not overlap with *SYDWG*. The name 'Zhan-Mao' area comes from the two areas in western Guangdong, Zhanjiang and Maoming.

Figure 12. Survey points in *SYDZM*

(14) Zhanjiang (Chikan), Lianjiang (Liancheng), Wuchuan (Wuyang), Suixi (Beipo), Maoming (Maonan), Gaozhou (Panzhoujie), Xinyi (Dongzhen), Dianbai (Yangjiao), Huazhou (Changqi)

It should be noted that in the Zhan-Mao area, Hakka and Min dialects are also spoken[32].

[32] See Figure 3 in Section 2.3

6.2.5 *The Lexicon of Chinese Dialects* (*Lexicon* hereafter)

The *Lexicon* (Beijing University Research Institute of Language 1989) is a database with phonetic transcriptions of 3,000 Chinese characters in 20 Sinitic varieties (8 Mandarin/Guanhua dialects, 2 Wu dialects, 2 Xiang dialects, 1 Gan dialect, 1 Hakka dialect, 2 Yue dialects and 4 Min dialects). The format of the *Lexicon* is also very similar to the Yue dialect surveys. The two Yue dialects are from Guangzhou and Yangjiang. Only the data for the Yangjiang dialect in *Lexicon* is used in this dissertation. There is a big gap between *SDPRD* and *SYDWG* where no data is available. The *Lexicon* provides one locality for this blank region. The location of Yangjiang is shown in the map with all other localities in Figure 13.

6.2.6 Localities

In total, the data of 54 dialects from the dialect surveys introduced above are used. The locations of all survey points are shown in Figure 13.

Figure 13. All the survey points investigated in this dissertation

6.3 Organisation/ layout of a typical Chinese dialect survey

It has been mentioned in Section 3.1.2 that Chinese dialectologists use the *Wordlist* (CASS 1983) to collect pronunciation data of Chinese characters from the informants. Since this wordlist is organized in the order of the rhyme categories in *Guangyun*, the dialect surveys present the data in the same order as the *wordlist*, for the ease of researchers to look up particular items. The dialect data (phonetic transcription of Chinese characters) are presented in table form.

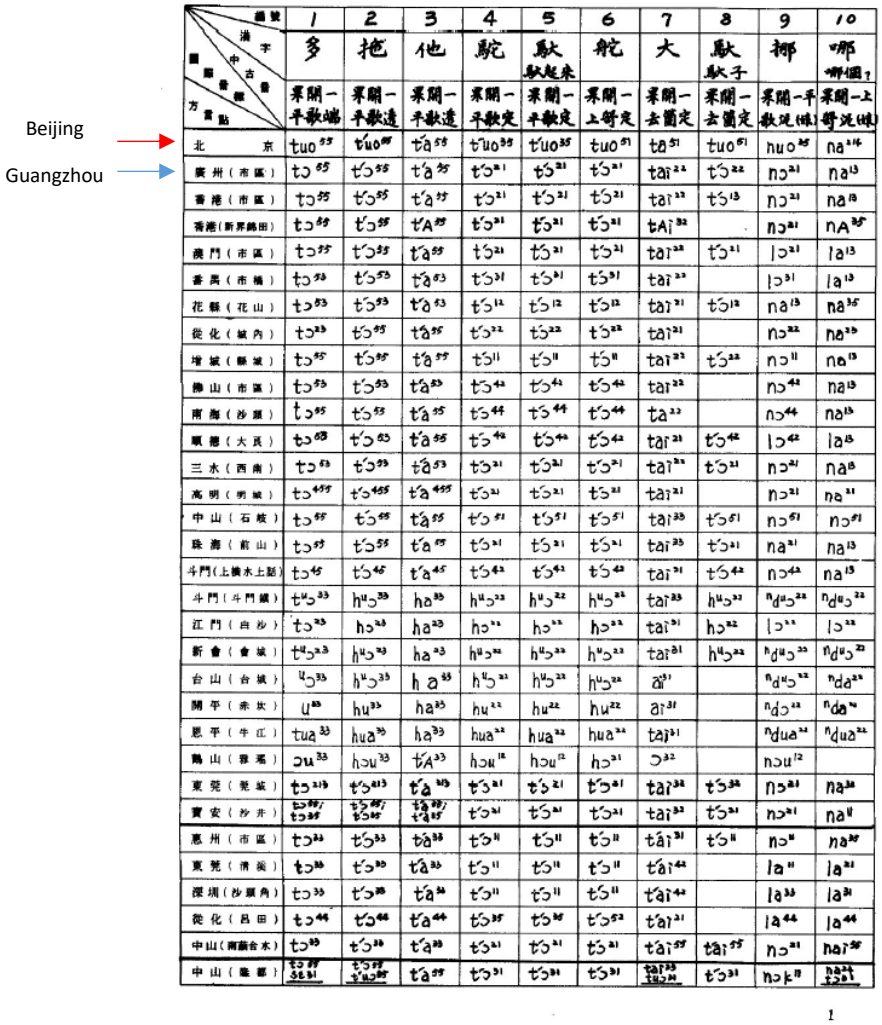

Figure 14. A page from *SDPRD* (Zhan & Cheung 1987: 1)

Figure 14 shows what a typical page of a dialect survey looks like. The leftmost column shows all the localities in the dialect survey. The Beijing dialect, like in Figure 14, is added to the list

for comparative purposes. The Beijing dialect (indicated with a red arrow) is usually at the very top row of the tableau. The Guangzhou dialect is indicated with a blue arrow. For other Yue dialect surveys, both Beijing and Guangzhou pronunciations are present for comparison.

The Chinese characters are numbered, as shown in the top row in Figure 14. The pronunciation of each character is transcribed in modified IPA for each dialect. This can be seen under each character. The most common modification of IPA used in Chinese dialectology is the symbol for aspiration. [’] is used for [^{h}] and it does not represent ejectives as in standard IPA (CASS 1983: 81). Additional vowel symbols, [ɿ], [ᴀ] and [ᴇ], can also be found. These symbols represent [ʐ̩], roughly [a̱] and [e̞]/[ɛ̞] (between [e] and [ɛ]) respectively (Handel 2015, Li 2017: 31).

In terms of tones, the tone contour is marked with numbers 1 to 5, which represent relative pitch. [^{1}] represents the lowest pitch and [^{5}] represents the highest (CASS 1983: 82). Level tones are marked with two identical numbers, e.g. [^{11}], which is a low level tone. A rising tone comprises a set of numbers from low to high and a falling tone is the opposite, e.g. [^{13}] and [^{42}] respectively. Lastly, more complex contours such as a HLH or LHL are represented with 3 numbers, e.g. [^{514}] and [^{141}] respectively (CASS 1983: 82). Checked syllables (syllable with a stop coda) are usually marked with one number.

The most complicated elements in the grid are the boxes in the third row from the top. Chinese dialectologists call this row ‘Middle Chinese’. They contain historical phonological information of the characters. Table 8 explains what information are present in the ‘Middle Chinese’ row[33].

[33] See Section 3.2 for the explanation of these categories.

Rhyme Group	Open/closed rhyme	Vowel Grade
Tone category	Rhyme category	Initial

Table 8. Middle Chinese sound categories

Since the *wordlist* is based on *Guangyun* (CASS 1983: v), the Middle Chinese sound categories from *Guangyun* are provided for each character. To find lexical items with a particular phonological feature, the researcher has to be familiar with the Middle Chinese sound categories in order to find all the items they need.

6.4 Data extraction

This dissertation is interested in the variation of present-day reflexes of MC *ɣu- in Yue dialects. The relevant information are already listed in Table 4 (Section 3.2.4), which I present here again:

Name (English)	**Does it have *closed rhymes*?**	*Ru* category?	**Vowel grades:**	**Can *ɣu- combine with this rhyme group (Grade I and II)?**
Guo	yes	no	I, III	yes
Jia	yes	no	II, III	yes
Yu	yes (no *open rhymes*)	no	I, III	yes
Xie	yes	no	I, II, III IV	yes
Zhi	yes	no	III	no
Xiao	no	no	I, II, III, IV	no
Liu	no	no	I, III	no
Xian	yes	yes	I, II, III, IV	no
Shen	yes	yes	I, II, III, IV	no
Shan	yes	yes	I, II, III, IV	yes
Zhen	yes	yes	I, III	yes
Dang	yes	yes	I, III	yes
Jiang	no	yes	II	no
Zeng	yes	yes	I, III	yes
Geng	yes	yes	II, III, IV	yes
Tong	yes (no *open rhymes*)	yes	I, III	yes

Table 4. List of Rhyme groups (from Section 3.2.4)

I have bolded the key information I need to consider before extracting the data. The MC *ɣu- sequence only appears in closed rhymes. Therefore, I only need to focus on rhyme groups that have closed rhymes. Next, it is worth mentioning again that this dissertation focuses on Grade I and II rhymes only (see Section 3.2.4). I leave out any rhymes that belong to Grade III and IV. Lastly, not all closed rhymes allow the combination with MC initial *ɣ-. I only focus on rhyme groups which a closed rhyme can combine with *ɣ-.

Taking consideration with the conditions mentioned above, only 10 rhyme groups fit these requirements for this dissertation. These rhyme groups consists of just over 70 words per dialect (the actual number fluctuates because not every single dialect has the same amount of data collected in the survey). They are listed in Table 9 below with MC reconstruction and pronunciation in Guangzhou Yue and Beijing Mandarin for reference.

There are possible instances of *field-worker isoglosses* (Trudgill 1983: 38) in the data. For example, some transcriptions recorded <w-> while others with <u->. Both segments are perceptually difficult to distinguish, especially through auditory transcription (no recordings were made in these surveys). Therefore, all the onset vowels are treated as a semivowel in my dataset.

Character	Gloss	Rhyme group	MC Vowel Grade	Middle Chinese	Guangzhou Yue	Beijing Mandarin
和 1	polite	Guo	I	*ɣuɑ	[wɔ]	[xɤ]
禾	grain	Guo	I	*ɣuɑ	[wɔ]	[xɤ]
禍	disaster	Guo	I	*ɣuɑ	[wɔ]	[xuo]
和 2	Huo noodle	Guo	I	*ɣuɑ	[wɔ]	[xuo]
華 1	Chinese	Jia	II	*ɣua	[wa]	[xua]
鏵	spade	Jia	II	*ɣua	[wa]	[xua]
劃	to row	Jia	II	*ɣua	[wa]	[xua]
華 2	Hua (name)	Jia	II	*ɣua	[wa]	[xua]
樺	birch	Jia	II	*ɣua	[wa]	[xua]
胡	Hu (name)	Yu	I	*ɣo/*ɣu	[wu]	[xu]
湖	lake	Yu	I	*ɣo/*ɣu	[wu]	[xu]
狐	fox	Yu	I	*ɣo/*ɣu	[wu]	[xu]
壺	pot	Yu	I	*ɣo/*ɣu	[wu]	[xu]
乎	(interrogative) particle	Yu	I	*ɣo/*ɣu	[fu]	[xu]
瓠	gourd	Yu	I	*ɣo/*ɣu	[wu]	[xu]
鬍	beard	Yu	I	*ɣo/*ɣu	[wu]	[xu]
戶	household	Yu	I	*ɣo/*ɣu	[wu]	[xu]
滬	Shanghai	Yu	I	*ɣo/*ɣu	[wu]	[xu]
互	mutual	Yu	I	*ɣo/*ɣu	[wu]	[xu]
護	to protect	Yu	I	*ɣo/*ɣu	[wu]	[xu]
回	to return	Xie	I	*ɣuɑi	[wui]	[xuei]
茴	fennel	Xie	I	*ɣuɑi	[wui]	[xuei]
匯	to exchange/ converge	Xie	I	*ɣuɑi	[wui]	[xuei]
潰	to rot	Xie	I	*ɣuɑi	[kʰui]	[xuei]
會 1	meeting	Xie	I	*ɣuɒi	[wui]	[xuei]
會 2	be able to	Xie	I	*ɣuɒi	[wui]	[xuei]
繪	to paint	Xie	I	*ɣuɒi	[kʰui]	[xuei]

Table 9. List of words analysed in this dissertation

Character	Gloss	Rhyme group	MC Vowel Grade	Middle Chinese	Guangzhou Yue	Beijing Mandarin
懷	to mind	Xie	II	*ɣuai	[wai]	[xuai]
槐	Pagoda Tree	Xie	II	*ɣuai	[wai]	[xuai]
淮	Huai (name)	Xie	II	*ɣuai	[wai]	[xuai]
壞	bad	Xie	II	*ɣuai	[wai]	[xuai]
畫	to draw/ drawing (n.)	Xie	II	*ɣuæ	[wa]	[xua]
話	speech/ to say	Xie	II	*ɣuɐi	[wa]	[xua]
桓	pillar (archaic)	Shan	I	*ɣuɑn	[wun]	[xuan]
完	to finish	Shan	I	*ɣuɑn	[jyn]	[uan]
丸	meat ball	Shan	I	*ɣuɑn	[jyn]	[uan]
緩	slow	Shan	I	*ɣuɑn	[wun]	[xuan]
皖	Anhui	Shan	I	*ɣuɑn	[wun]	[uan]
換	to change/ switch	Shan	I	*ɣuɑn	[wun]	[xuan]
活	to live/ alive	Shan	I	*ɣuɑt	[wut]	[xuo]
幻	fantasy/ magical	Shan	II	*ɣuɐn	[wan]	[xuan]
滑	slippery	Shan	II	*ɣuɐt	[wat]	[xua]
猾	sly	Shan	II	*ɣuɐt	[wat]	[xua]
還 1	to return/ revert	Shan	II	*ɣuan	[wan]	[xuan]
還 2	besides	Shan	II	*ɣuan	[wan]	[xai]
環	ring	Shan	II	*ɣuan	[wan]	[xuan]
患	to suffer from	Shan	II	*ɣuan	[wan]	[xuan]
宦	eunuch	Shan	II	*ɣuan	[wan]	[xuan]
魂	spirit	Zhen	I	*ɣuən	[wɐn]	[xuən]
餛	'won' in wonton	Zhen	I	*ɣuən	[wɐn]	[xuən]
渾	muddy	Zhen	I	*ɣuən	[wɐn]	[xuən]
混	to mix	Zhen	I	*ɣuən	[wɐn]	[xuən]
核	core	Zhen	I	*ɣuət	[wɐt]	[xu]

Table 9. (cont.)

Character	Gloss	Rhyme group	MC Vowel Grade	Middle Chinese	Guangzhou Yue	Beijing Mandarin
黃	yellow	Dang	I	*ɣuɑŋ	[wɔŋ]	[xuaŋ]
簧	reed	Dang	I	*ɣuɑŋ	[wɔŋ]	[xuaŋ]
皇	king	Dang	I	*ɣuɑŋ	[wɔŋ]	[xuaŋ]
蝗	locus	Dang	I	*ɣuɑŋ	[wɔŋ]	[xuaŋ]
晃	to sway	Dang	I	*ɣuɑŋ	[fɔŋ]	[xuaŋ]
鑊	wok	Dang	I	*ɣuɑk	[wɔk]	[xuo]
弘	great	Zeng	I	*ɣuəŋ	[wɐŋ]	[xuŋ]
或	or	Zeng	I	*ɣuək	[wak]	[xuo]
惑	confused	Zeng	I	*ɣuək	[wak]	[xuo]
橫 1	horizontal	Geng	II	*ɣuɐŋ	[waŋ]	[xəŋ]
橫 2	peremptory	Geng	II	*ɣuɐŋ	[waŋ]	[xəŋ]
宏	grand	Geng	II	*ɣuaŋ	[wɐŋ]	[xuŋ]
獲	to gain	Geng	II	*ɣuak	[wɔk]	[xuo]
劃	to draw	Geng	II	*ɣuak	[wak]	[xua]
鴻	red	Tong	I	*ɣuŋ	[hoŋ]	[xuŋ]
紅	vast/ flood	Tong	I	*ɣuŋ	[hoŋ]	[xuŋ]
洪	large	Tong	I	*ɣuŋ	[hoŋ]	[xuŋ]
虹	rainbow	Tong	I	*ɣuŋ	[hoŋ]	[xuŋ]
汞	mercury	Tong	I	*ɣuŋ	[hoŋ]	[kuŋ]
鬨	to coax	Tong	I	*ɣuŋ	[hoŋ]	[xuŋ]

Table 9. (cont.)

7. Analysis

In this section, I first describe how the phonological environments of the lexical items are analysed in the analysis. Then, I first give a general inspection and classification of the dialect reflex patterns, followed by a dialectometrical analysis to validate my inspection. Each dialect pattern is then analysed in detail.

7.1 The analysis of the phonological environments

The phonological environments analysed in this dissertation are based on the phonetic realisations found in the dialect surveys, not the Middle Chinese rhyme group categories. I make the assumption that what we see in the dialect survey reflects the earlier stage of the dialect when ɣ-deletion and/or ɣ-devoicing took place. For example, the [f-] reflex is found in the pre-u# environments in the Nanhai dialect. In the survey, both *Yu* *u/*o and *Xie* *uɑi rhyme groups are reflected as [u] in the Nanhai dialect. My assumption is that when ɣ-devoicing took place[34] in this dialect, the phonological environment for this change would be a *u, not both MC *u/*o and *uɑi as the rhyme group categories suggested. This assumption allows me to analyse the dialect-specific sound changes for which evidence is only present in the survey data. This assumption is only disregarded if there is enough evidence for an alternate phonological environment at an earlier stage of the variety.

7.2 Patterns found in dialects

From the previous descriptions of the historical changes of MC *ɣu-, the Guangzhou and Dongguan dialects have provided a typology of reflex patterns for Yue dialects (see Figure 8 in Section 5.2.5). Just as a quick recap, the Guangzhou dialect had ɣ-loss before the MC medial *-u- and nucleus *u (< *o). The Dongguan dialects and the rural Guangzhou dialects show another pattern: ɣ-loss is found in most phonological environments (reflected as [w-] in rural Guangzhou dialects and [v-] in the Dongguan dialects), except for pre-u# contexts, where the

[34] See Figure 8 in Section 5.2.5.

reflex [f-] is found, as a result of ɣ-devoicing and labio-dentalisation.

After eye-balling the whole dataset, I have found more dialect patterns than the two mentioned above. The general patterns are summarized in Table 10. The sound change typology comes from Figure 8 in Section 5.2.5: Type 1 and 2 changes refer to ɣ-deletion and glide-insertion; Type 3 and 4 changes are related to ɣ-devoicing.

Representative dialect	**Sound changes**	**Traditional Yue dialect classification**	
		Zhan (2002)	***LAC*[35] (2012)**
Guangzhou	Type 1	Yuehai, Xijiang	Guangfu, Goulou
Taishan	Type 2a	Siyi, Yuehai	Siyi, Guangfu
Dongguan (Guancheng)	Type 2a & Type 3 (Phon. Conditioned)	Guan-Bao	Guangfu
Bao'an	Type 1 & Type 3 (Phon. Conditioned)	Guan-Bao	Guangfu
Maoming	Type 2b & Type 3 (Phon. Conditioned)	Gao-Lei	Gao-Yang
Zhongshan	Type 1 & Type 4 (Phon. Conditioned)	Xiangshan	Guangfu
Lianjiang	Type 1 & Type 2b, 3 (Phon. Conditioned)	Gao-Lei	Gao-Yang
Gaozhao	Type 1, 2b & Type 3 (Phon. Conditioned)	Gao-Lei	Gao-Yang

Table 10. Summary of dialect patterns according to their reflexes of MC *ɣu-

Based on Table 10, there could be 8 types of dialects in Guangdong. It turns out that there are dialects in which Type 1 (ɣ-deletion) and Type 2a (w-fortition) changes could occur across all the phonological environments, but not for the other types. Other changes such as Type 3 (ɣ-devoicing followed by x-labio-dentalisation) are usually phonologically conditioned,

[35] *Language Atlas of China*

accompanied by another change. There is no dialect that has the reflex [f-] (Type 3) across all environments. However, in some cases, like in the Gaozhao dialect, the phonological distribution of [w-] seems to be more unpredictable. The distribution of the reflexes of MC *ɣu- is analysed in more detail in Section 7.4 and is discussed in Section 8.

These groupings do not quite correspond to the traditional classifications (Zhan 2002, CASS & CityU LISRC 2012: 129) of these dialects. For example, within the Yuehai, Guan-Bao and Gao-Yang dialect groups, there are different dialect patterns, see Table 10. Furthermore, the Huaxian dialect is traditionally classified as a Guangfu dialect (Xiong 1987: 161, LAC 2012), but it actually patterns with the Siyi dialects, despite geographical distance between the Huaxian dialect and the Siyi dialects.

Eye-balling the dialect data is not an accurate way to identify dialect patterns based on their reflexes, as this method is subject to human error. The generalizations I made in Table 10 also left out some details, e.g. possible lexical exceptions and the phonological contexts for [f-]. The grouping of dialect patterns requires a more quantitative and objective method.

7.3 Dialectometry and Multidimensional Scaling

To seek clusters that are similar in terms of their linguistic structure, I need dialectometry. Dialectometry can convert qualitative data (data found in dialect atlases and surveys) to quantitative data without reference to geography. There are two main approaches: (a) The Salzburg approach: essentially converting linguistic atlas data into measures of similarity between dialects, which is shown in Figure 15 (Goebl 2010b: 437) and (b) The Groningen approach: the transcriptions from the linguistic atlas for each dialect are compared with each other, segment to segment (with the Levenshtein algorithm, see Heeringas (2004), Nerbonne & Heeringa (2010: 550-567) for more details) to produce a quantified dialect distance.

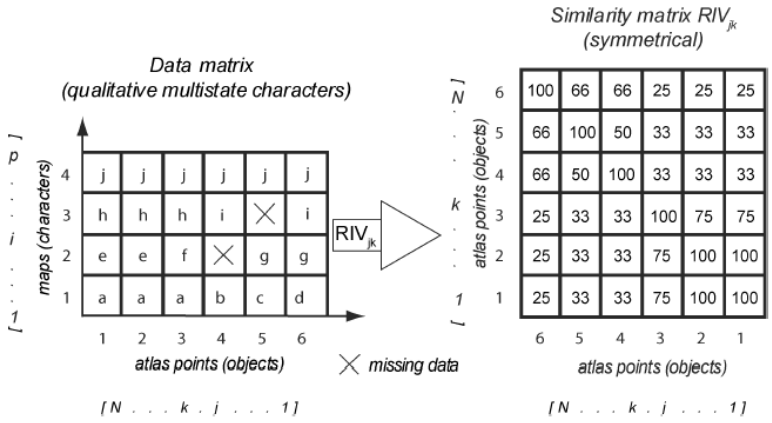

Figure 15. From data matrix to similarity matrix (Goebl 2010b: 437)

The Salzburg approach is suitable for this dissertation, since I am dealing with categorical data. I am only comparing the reflexes of the initial in each word amongst all the dialects, and not comparing different pronunciation of words that contains more than one segment.

Figure 15 (taken from Goebl 2010b: 437) illustrates how to convert qualitative data, i.e. the reflexes of MC *ɣu- in each dialect, to dialect similarity. The grid on the left in Figure 15 represents the matrix of categorical data that I extracted from the dialect surveys. The rows represent the words and the columns represent the localities. Each cell shows the corresponding reflex found in a particular word in the relevant locality. An algorithm, *Relative Identity Value* (RIV), is used to produce the similarity matrix (on the right in Figure 15). RIV is presented in the formula in (15) below (Goebl 2010b: 439).

(15) $$RIV_{jk} = 100 \sum COI(i)_{jk} / \sum COI(i)_{jk} + \sum COD(i)_{jk}$$

RIV is a similarity index which calculates the proportion of shared items out of all the items compared between two dialects (Goebl 2010b: 439). (15) shows that to calculate the RIV of *j* and *k* (two dialects), one has to divide the number of items which is shared by both dialects (COI, co-identity) by the total number of items (COI & COD, Co-difference) between these two dialects. Hence, it can be reformulated as in (16) for readers who are not as familiar with

the computational formulation.

(16) $RIV_{Dialect\ a\ \&\ Dialect\ b}$ = shared features / total number of features compared x 100%

Using the formula in (15) to calculate the RIV for all dialect pairs, a similarity matrix can be produced, as in the grid on the right in Figure 15. This matrix shows the similarity between all the dialect pairs. The interpreter only has to read half of the matrix, since the opposite half is exactly identical. To find the similarity between, e.g. dialect *a* and *b*, one firstly finds the column for dialect *a* and then the row for dialect *b*. The intersecting cell for this column and row is the RIV between dialect *a* and *b*. In this dissertation, a distance matrix is calculated instead because the next procedure below requires a distance matrix. The calculation for the distance matrix is 100-RIV (reformulated from Goebl 2010b: 440).

The distance matrix provides the linguistic distance for each dialect pair, but it is uninterpretable to human eyes and therefore requires visualization. This analysis employs multidimensional scaling for this purpose. Multidimensional Scaling (MDS) is a "statistical technique aimed at representing very high dimensional data in a smaller number of dimensions" (Nerbonne 2010a: 467). In other words, MDS converts the distances from a distance matrix into a (e.g. 2-dimensional) plot computationally (Embleton 1993, Heeringa 2004: 156, Nerbonne & Wieling 2018: 403), so that the analyst can interpret the information easily. In this plot, linguistic distances are roughly represented in Euclidean distance on a plane. The higher the correlation[36] of the distance in this plot with the distances of the elements in the matrix, the more successful the 'dimension reduction' is (Nerbonne & Wieling 2018: 403).

[36] Indicated by the the correlation coefficient *r* (Nerbonne & Wieling 2018: 403).

Figure 16 is an MDS plot produced with *Gabmap* (Nerbonne et al. 2011: 65-89; Leionen 2016: 71-83). The dataset is checked to be very consistent[37] (Cronbach's α = 0.99). The correlation is very high (r = 0.99), it can explain 98% (r^2) of the data in the distance matrix.

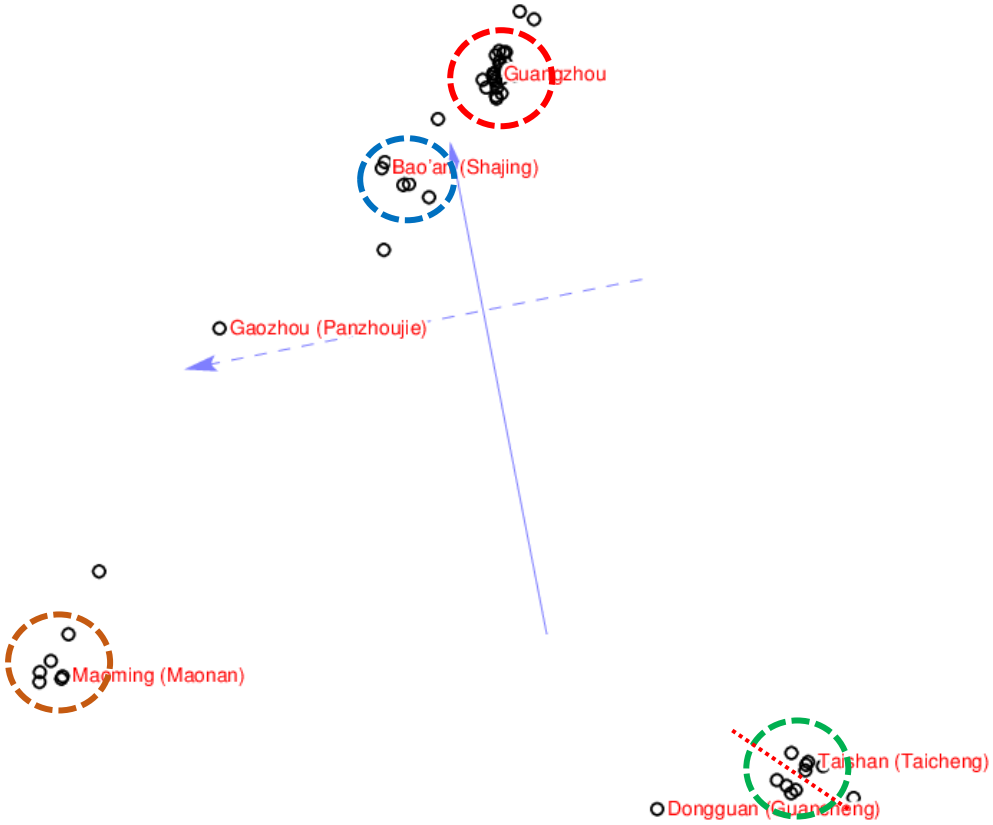

Figure 16. MDS plot of the dialect distances

Big distances between dialects and clusters in the MDS plot reflect big linguistic distances between them. In Figure 16, there are three very distinct clusters and a number of outliers. Moreover, two more clusters can be separated from the biggest group, namely clusters circled in red and blue.

[37] Cronbach's α is a method that measures consistency or reliability of the data (for the dialectometrical analysis). Consistency/reliability is shown with the coefficient α; the value is between 0 and 1, with 1 being very consistent/reliable. See more in Heeringa & Prokic (2018: 340).

The tightest and densest cluster is where Guangzhou is located in the plot, circled in red. I call these dialects the Guangzhou-type dialects. A looser cluster is located near Guangzhou-type dialects, namely the cluster near Bao'an. Dialects in this cluster are called Bao'an-type dialects, circled in blue. The fact that the Bao'an cluster is situated much closer to the Guangzhou cluster than the other two clusters in the MDS plot suggests that these two clusters are much more similar to each other than the rest of the clusters. Another cluster is found near the Taishan dialect, which also has a lot of members. I name dialects in this cluster the Taishan-type dialects, circled in green. A loose cluster with fewer members lies near Maoming (in the brown circle). I name these dialects Maoming-type dialects. Other than these four clusters, there are dialects located farther away from the clusters (e.g. Gaozhou between the Bao'an and Maoming clusters). They are analysed in Section 7.4.5 separately from the other clusters as outliers/transitional dialects.

The clusters in Figure 16 matches some of the observations I made about the dialect patterns in Table 10. However, there are only 4 clusters instead of the 8 groups which I observed. There are a few possibilities as to why this is the case. Perhaps the generalisations I made in Table 10 are too narrow, producing a large number of patterns, or rather, the way I distinguish clusters in Figure 16 is too naïve and subjective – maybe I have ignored internal divisions within a 'cluster' (coloured circles). For example, within the green dotted circle, elements can perhaps be further divided into top and bottom half, as shown with the dotted red line in Figure 16. The outliers may also form a pattern of their own or it could belong to the closest cluster. What these clusters in the MDS plot actually stand for is explored in Section 7.4.

It should be recognised that the application of dialectometry and MDS here is not to classify dialects within Yue. Using a (random) single feature to classify dialects can result dialect groupings, but it is not known whether this feature is truly significant for geographical

differences or not (Nerbonne 2010a: 477). MDS is used here simply to visualize and identify dialects which share the same dialect pattern so that dialects can be described more systematically. MDS helps me to validate my observations in Table 10 with a more objective method before explaining why such dialects have the current patterns.

As Table 10 demonstrated, the traditional classification is not very successful at grouping dialects together in terms of their reflexes of MC *ɣu-. Zhan (2002, see Figure 7) and *LAC* (2012) all show more than 4 sub-dialect groups, each with different members from what the MDS plot shows. I therefore will not explore the dialect patterns according to their traditional classifications. The following subsections in Section 7.4 are guided by the main clusters found in Figure 16.

7.4 Dialect clusters and their reflexes of MC *ɣu-

The MDS plot in Section 7.3 has shown four clusters, suggesting there could be four major dialect pattern across Yue dialects in Guangdong. The four clusters are analysed in this order: Guangzhou-, Taishan-, Maoming-, Bao'an-type. The outliers in the MDS plot are analysed in a separate subsection.

In each subsection below, a cluster from above is analysed in detail – the phonological pattern of the representative dialect in each cluster is firstly described, and outliers of the described features are identified. Next, other dialects of the same cluster are listed and their features are also described. Outliers of their phonological patterns are pointed out afterwards. Lastly, 'words' and '(lexical) items' are used interchangeably here, since the data consists of monosyllables.

Figure 8 (Section 5.2.5) is the guideline for identifying words that went through either ɣ-deletion (Type 1, 2a, 2b) or ɣ-devoicing (Type 3, 4). In Section 7.2, I have shown that ɣ-deletion is prevalent in all dialects. In order to assess the regularity of ɣ-deletion, reflexes other than [w-], [v-], [ʋ-] as well as the ones not stated in Figure 8, e.g. [k-] and [j-], are identified (labelled as 'possible exceptions') and assessed. Borrowings, contact phenomena, archaisms/remnants from previous strata[38] and phonologically conditioned ɣ-devoicing are not considered as true exceptions. Moreover, attentions are paid to the reflex [f-]. Li (1996) proposes that the reflex [f-] is an exception to the change whereas Lau (2003) believes that forms with the reflex [f-] are borrowings. I aim to address the origin of the reflex [f-] in Section 8.2.

There are items which seem to be general lexical exceptions, i.e. items that appear in the same or similar forms regardless of the dialect cluster. These items are listed in Table 11. These items are temporarily taken out from the analysis to avoid repetition for all clusters. They are all addressed in the discussion in Section 8.1 with a discussion on regularity of sound changes.

Reflex	Word	Gloss	Reflex	Word	Gloss
f	乎	Interrogative particle	h/f	宏	grand
f	晃	to sway	h/k	紅	red
j/z	完	to finish	h/k	洪	flood
j/z	丸	meat ball	h/k	鴻	large
k^h/k	潰	to rot	h/k	虹	rainbow
k^h	繪	to paint	h/k/k^h	汞	mercury
h/f	弘	great	h/k	鬨	coax

Table 11. General lexical exceptions

7.4.1 Guangzhou-type dialects (red circle)

Other than the general lexical exceptions to ɣ-deletion listed in Table 11, the Guangzhou dialect shows no other words that show evidence of ɣ-devoicing. The remaining items are found to

[38] In Section 4.3.2, I have introduced Mai's (2009) concept of strata-swamping. 'Archaisms/remnants' here refers to the forms/retentions that were not replaced by the newer stratum after swamping; these reflexes could have taken part in newer sound changes too, resulting forms that neither resemble ɣ-deletion nor ɣ-devoicing.

begin with [w-].

Other Guangzhou-type dialects include:
Congua, Deqing, Fogang, Foshan, Gaoming, Guangning, Hong Kong (Urban), Huaiji, Lechang, Lianxian, Luoding, Macau, Panyu, Qingyuan, Qujian, Renhua, Sanshui, Shaoguan, Sihui, Yangjiang, Yangshan, Yunan, Yunfu, Zhongshan (24 dialects in total).

The dialects listed above form the biggest cluster in the whole data on the MDS plot. Geographically speaking, they are distributed around the Pearl River Delta as well as Northern and Central-western Guangdong (up to Luoding and Yangjiang) with the exceptions of the Siyi area.

These dialects pattern like Guangzhou in that most of the reflexes of MC *ɣu- are [w-]. Reflexes other than [w-] are listed in Table 12 followed by an analysis.

Dialect	Character	Gloss	Pronunciation
Conghua, Fogang, Lianxian^{2}, Zhongshan^{3}	桓	pillar (archaic)	fun, jyn^{2}, hun^{3}
Conghua	換	to change/ switch	fun
Yangjiang, Zhongshan^{2}	互	mutual	fu, hu^{2}
Yangjiang, Zhongshan^{2}	護	to protect	fu, hu^{2}
Macau, Qujiang, Zhongshan^{2}	瓠	gourd	fu, hu^{2}
Macau, Qujiang, Zhongshan^{2}	鬍	beard	fu, hu^{2}
Renhua, Zhongshan^{2}	滬	Shanghai	fu, hu^{2}
Gaoming, Lianxian, Macau, Qingyuan, Qujiang, Renhua, Sanshui, Yunfu, Zhongshan	核	core	hɐt
Gaoming, Hong Kong (urban), Lechang, Qingyuan, Qujiang, Renhua, Shaoguan, Yunfu	皖	Anhui (place name)	jyn
Qujiang, Yangjiang^{2}, Zhongshan^{3}	緩	slow	jyn, fun^{2}, hyn^{3}
Gaoming, Sanshui^{2}	禾	grain	ho, fɔ^{2}

Gaoming	和 1	polite	ho
Gaoming	和 2	Huo noodle	ho
Gaoming	黃	yellow	huoŋ
Gaoming	簧	reed	huoŋ
Gaoming	皇	king	huoŋ
Gaoming	蝗	locus	huoŋ
Guangning, Huaji, Zhongshan[2]	狐	fox	ku, hu[2]
Guangning	槐	Pagoda tree	kuai
Yangjiang	混	to mix	kuɐn
Macau	餛	'won' in 'wonton'	$k^huɐn$
Huaiji	劃	to row	pa
Sanshui	樺	birch	fa
Sanshui	鑊	wok	kuɔk
Yangjiang	渾	muddy	fɐn
Zhongshan	胡	Hu (name)	hu
Zhongshan	湖	lake	hu
Zhongshan	壺	pot	hu
Zhongshan	戶	household	hu
Zhongshan	回	return	hu
Zhongshan	茴	fennel	hu

Table 12. Possible exceptions to ɣ-deletion in other Guangzhou-type dialects

Table 12 lists all the possible exceptions to ɣ-deletion, i.e. syllables which do not have an initial [w-], in the Guangzhou-type dialects. A word can have different pronunciations in different dialects. If the location has a superscripted number, the pronunciation of the word in that dialect is indicated by the same superscripted number next to it.

Figure 17 shows the variation of reflexes for 'core' in Guangdong. The pronunciation for 'core' shows no distinct geolinguistic areas for these reflexes. [w-] and [h-] are the most common reflexes for this word. ɣ-deletion is more commonly found in western and northern Guangdong (purple, orange circles); localities with two pronunciations are found mainly in the Siyi area (light blue) and its surrounding area (pink); ɣ-devoicing is found in the Siyi area and northern

Guangdong (red circles) and south-eastern Pearl River Delta (green). According to the dialect surveys (Zhan & Cheung 1987, Zhan & Cheung 1994), [w-] and [h-] are *colloquial* and *literary pronunciations* [39] (Li 2007: 93). The colloquial pronunciation reflects the inherited pronunciation of the character and the literary pronunciation reflects the borrowed pronunciation (Li 2007: 93). In the Renhua dialect, [wɐt] is indicated as the colloquial pronunciation (Zhan & Cheung 1994: 350). If we accept their indication, then ɣ-deletion would be an endogenous change for Yue, which is very plausible as it is an innovation found across all Yue dialects (mentioned at the start of Section 7.4). It also suggests the [hɐt] pronunciation is borrowed. The origin of [hɐt] is currently unknown, presumably it was from a historical prestigious variety. If we have accept that the [hɐt] pronunciation was borrowed, it predicts that the localities that only have [hɐt] have lost their native pronunciation for this word.

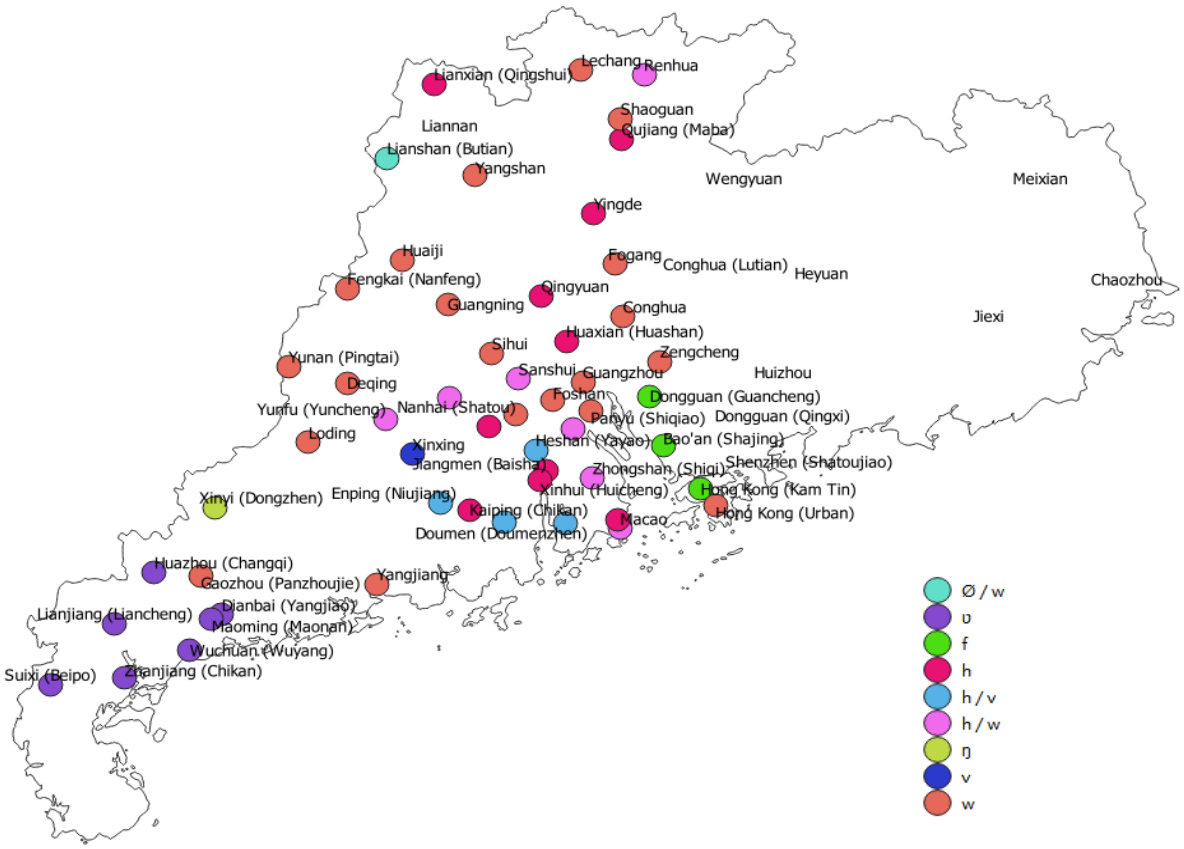

Figure 17. Point-colour map of ‘core’

[39] See also Section 5.2.3.

There are words which are pronounced as [jyn], such as ‘pillar’, ‘Anhui’ and ‘slow’. I treat these words as the same as ‘to finish’ and ‘meat ball’ in Table 11 since they share the same form. See Section 8.1 for their discussion.

There are a number of words that begin with [f-] as a reflex of MC *ɣu-, namely ‘pillar’, ‘to change/switch’, ‘mutual’, ‘to protect’, ‘gourd’, ‘beard’, ‘Shanghai’, ‘great’, ‘grand’ and ‘grain’. ‘Pillar’ and ‘to change/switch’ are pronounced as [fun] and the rest [fu]. ‘Pillar’ is an archaic word that is used in idioms. This pronunciation could be borrowed from Guanhua in the past. It is unclear why ‘to change/switch’ has the [fun] pronunciation. For the rest of the words, [fu] is found in the Yangjiang, Macau, Qujiang and Renhua. Yangjiang is situated next to the Zhan-Mao dialects, where [fu] is found. This suggests that Yangjiang is related to these dialects in some ways, which I will come back in the later sections with more evidence. The Macau dialect used to be closely related to the Xiangshan dialects (Zhan 2002), in which the [fu] pronunciation is found. The relationship between Macau and the Xiangshan dialects is discussed in Section 8.2 and it explains why [fu] is found in Macau. Lastly, Qujiang and Renhua are both located in northern Guangdong, surrounded by the Hakka-speaking region. A possible explanation for the [fu] pronunciation is that they are borrowings from Hakka. Lastly, ‘grain’ and ‘birch’ are pronounced as [fɔ] and [fa] respectively in Sanshui. The possible reason behind this is discussed in Section 8.2.

There are instances of the reflex [k-] found in ‘fox’, ‘pagoda tree’ and ‘to mix’. They are mainly found in dialects spoken in central-western Guangdong, just like Zhan’s (2002: 177) description of some Xijiang dialects (see Section 5.2.2). However, these few instances of [k-] are sporadic. Wan’s (1995) analysis could shed some light on the origin of these initials. Wan found [k^h-] initials in his Anyi dialect data. Middle Chinese *ɣ- came from Old Chinese **g-. Wan (1995: 234) believes that [k^h-] results from a handful of initials which failed to participate

in the split from **g- to *ɣ- in Middle Chinese, followed by MC obstruent devoicing and aspiration, i.e. OC **g- > MC *g- > k^h-. I suspect the [k-] reflexes could be local pre-*Qieyun* archaisms in these dialects too. Like Wan proposed, I believe that the [k-] initial followed a similar path, except the devoicing generated unaspirated stops. Xijiang/Goulou dialects are known for having unaspirated voiceless stop from MC obstruent devoicing (see Figure 5), which supports the idea that this initial could be from MC *g- (< OC **g-). This explains why we do not see [w-] as a reflex, since *g- was not a target of ɣ-deletion. For the rest of the unexplained [k-] initials in other dialects, I believe that they could be sporadic archaisms[40] too.

The Gaoming dialect shows a number of instances of reflex [h-]. These instances are largely phonologically conditioned (i.e. before [-uoŋ] and [-o]), therefore they should not be considered as exceptions to ɣ-deletion. However, there is one word, 'disaster' [wo], which looks like ɣ-deletion applied. There are two possibilities for this: 1) it is a borrowing from neighbouring dialects and 2) there was an incomplete change, leaving one instance of [wo] and others [ho]. Without more evidence, a conclusion cannot be made for this item.

I argue that the Zhongshan dialect does not belong to this cluster. The Zhongshan dialect is closely related to the Zhuhai dialect, which belongs to the Bao'an-type dialects. Thus, the phonological environments that the reflex [h-] occurs in match the Bao'an-type dialects. Therefore, the discussion of the Zhongshan dialect is resumed in the analysis of the Bao'an-type dialects in Section 7.4.4.

The findings in the Guangzhou-type cluster are:

1. ɣ-deletion is the prevalent sound change for these words; [w-] (Type 1) is found across the

[40] Just as a reminder, 'archaism' here refers to forms/retentions that were not swamped; they could participate in newer changes if they fit the conditions of the change.

Guangzhou-type dialects.

2. ‘core’ has two pronunciations in a number of localities; [wɐt] is the result of ɣ-deletion and [hɐt] is the borrowed pronunciation of this word.
3. [f-] is found in a number of dialects; they seem to come from contact with neighbouring dialects (e.g. Qujiang).
4. [k-] is indeed found in some central-western dialects as previously found, but there are only few instances of them and they are sporadic. These could be local pre-*Qieyun* archaisms undergoing MC obstruent devoicing.
5. The Gaoming dialect shows numerous instances of [h-], but they are mostly phonologically conditioned (by [-uoŋ] and [-o]). ‘Disaster’ [wo] is an exception for this generalization; its origin is currently unknown.
6. Zhongshan belongs to the Bao’an-type dialect based on its structural similarity.
7. Besides borrowings, archaisms and dialect contact, ɣ-deletion in most Guangzhou-type dialects shows regularity.

7.4.2 Taishan-type dialects (green circle)

The Taishan dialect is somewhat similar to the Guangzhou-type dialects, in terms of the possible existence of ɣ-deletion in its phonological history. In section 7.4.1, I show that the Guangzhou-type dialects generally have had ɣ-deletion and the main reflex of MC *ɣu- is [w-]. In the Taishan dialect, the reflex is mainly [v-] (Type 2a), which corresponds to Guangzhou [w-], like Zhan (2002: 146) previously described. Examples can be found in Table 13. If the phonological conditioning and exceptions of ɣ-deletion are similar to the Guangzhou-type dialects, it suggests these two groups of dialects might be closely-related in the development of MC *ɣu-. Based on the data, the environment that triggered ɣ-deletion is identical to the Guangzhou-type dialects, i.e. preceding MC *-u- or *-u.

Character	Gloss	Taishan	Guangzhou
華	Chinese	va	wa
湖	lake	vu	wu
換	to change	vᵘɔn	wun
環	ring	van	wan

Table 13. Sample words from the Taishan and Guangzhou dialects

Like the Guangzhou-type dialects, there are words which appear to be exceptions to ɣ-deletion at first glance. These items are listed below in Table 14.

Character	Gloss	Pronuncition
劃	to row	pʰa
瓠	gourd	fu
戶	household	fu
滬	Shanghai	fu
緩	slow	fᵘɔn
皖	Anhui	fᵘɔn
核	core	hʌt (Std. IPA [hạt])

Table 14. Possible exceptions of ɣ-deletion in the Taishan dialect.

Figure 18 shows various reflexes of *ɣu- for ‘to row’. The most abundant reflex is [w-] (light blue), found in most dialects in Guangdong. In the Zhan-Mao area, [ʊ-] is the prevalent reflex (red circles). [v-] (in magenta) is found in Heshan and Jiangmen dialects in the Siyi area as well as Dongguan, Huaxian and Yingde. Huaiji and Kaiping have [p] (yellow) and [h] (orange) reflex respectively. Lastly, most of the Siyi dialects have [pʰ] in the reflex of ‘to row’.

[pʰa] appears to be a region-specific variant for ‘to row’ in the Siyi area. This matches Yan’s (2006: 209-210) description, see (5) in Section 5.2.1 (Xinhui pronounces [pʰua], which I consider as a sub-variant of the form). I argue that [pʰ-] could be a local pre-*Qieyun* archaism for a number of reasons. Firstly, it does not look like an innovation as it would be unusual for *ɣu- to change to [pʰ-]. Secondly, there are other words, e.g. ‘Chinese’ *ɣua, which had the

exact same MC form, but the present-day realization is different (see Table 13). If it were an innovation, why did it only affect 'to row' but not other words? Since this form is region-restricted, the explanation could be that it is another local pre-*Qieyun* archaism. 'To row' is therefore not considered as an exception to ɣ-deletion here, since [p^h-] is not relevant in this sound change.

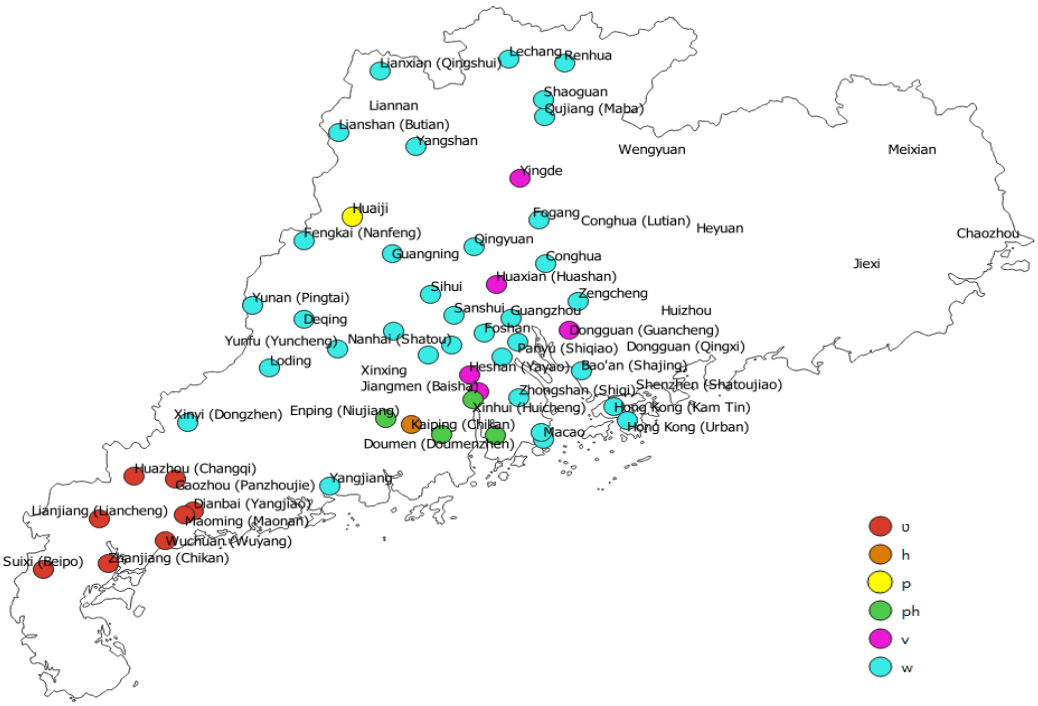

Figure 18. Point-colour map of 'to row'

For the pre-u# environment and before [-^{u}ɔn], instances of the [f-] reflex are found (see Table 14). In these environments, the majority of the reflexes are [v-], like 'lake' and 'to change' in Table 13. Therefore, the reflexes [f-] and [v-] are in parallel distribution. It seems like these words with [f-] are exceptions to ɣ-deletion, but I argue that the reflex [f-] are not true exceptions. I believe that reflex [f-] came from an older stratum which underwent ɣ-devoicing. This stratum was later swamped by the Guangzhou-type dialects (which already underwent ɣ-deletion by that time) and some words like 'slow' did not get replaced by the new stratum. I leave the arguments for this proposal to Section 8.2 after reviewing the data from other dialects.

The other Taishan-type dialects are: *Yingde, Heshan, Huanxian, Jiangmen, Doumen, Engping, Kaiping* and *Xinhui*. All of them are Siyi dialects except for Yingde and Huaxian, which are considered traditionally as Guangfu/Yuehai dialects (CASS & CityU LISRC 2012, Zhan 2002: 119). The reason for why they are clustered together seems to be that these two dialects also have [v-] as their main reflex for MC *ɣu-.

As before, all instances which appear to be exceptions to ɣ-deletion are listed below in Table 15 and a discussion follows as to whether they are true exceptions are not.

Dialect	Character	Gloss	Pronunciation
Heshan	禾	grain	jyɵ
Heshan	和	polite	jyɵ
Heshan	和	Huo noodle	jyɵ
Heshan	禍	disaster	jyɵ
Heshan, Doumen, Enping, Kaiping, Xinhui	桓	pillar	fun
Heshan^{2}, Jiangmen, Doumen, Enping^{2}, Kaiping^{2}, Xinhui^{3}, Huaxian, Yingde	核	core	hɐt, hat^{2}, hæt^{3}
Jiangmen, Doumen, Enping, Kaiping, Xinhui	戶	household	fu
Jiangmen, Doumen, Enping, Kaiping, Xinhui	滬	Shanghai	fu
Jiangmen, Doumen, Enping^{2}, Kaiping^{2}, Xinhui	緩	slow	fun, fuan^{2}
Jiangmen	鏵	Spade	p^{h}a
Doumen, Enping, Kaiping^{2}, Xinhui^{3}	劃	to row	p^{h}a, ha^{2}, p^{h}ua^{3}
Enping, Kaiping	皖	Anhui	fuan
Heshan^{2}, Jiangmen	樺	Birch	fa, f$_{A}^{2}$ (Std. IPA [fa̱])

Table 15. Possible exceptions to ɣ-deletion in other Taishan-type dialects

The words listed in Table 15 are very similar to the possible exceptions for the Taishan dialect, as well as the ones in the Guangzhou-type dialects in Table 12 (excluding the Gaoming-,

Zhongshan-specific words). This suggests a link between the Taishan-type dialects and the Guangzhou-type dialects that might be missing in the traditional classifications.

The reflex [f-] is found ‘household’, ‘Shanghai’, ‘slow’, ‘pillar’, ‘Anhui’ and ‘birch’. These items in which [f-] is found overlap with the ones found in the Taishan dialect. Moreover, they are also in parallel distribution with [v-] in these environments. This means that for these dialects, there also appears to be exceptions to ɣ-deletion. The occurrence of reflex [f-] in these dialects and the Taishan dialect is addressed in Section 8.2 along with its explanation.

An interesting observation is that only the Siyi dialects exhibit reflex [f-]. Huaxian and Yingde dialects (traditionally classified as Guangfu, roughly Guangzhou-type dialects) do not share this feature (and they do not have reflex [p^h-] in ‘to row’). It seems like Taishan-type dialects are grouped together by their abundance of [v-]; structurally, the Huaxian and Yingde dialects are much closer to Guangzhou-type dialects rather than the Siyi dialects. Therefore, they should be considered as outliers of the Guangzhou-type dialects.

It should be noted that in Table 15, the pronunciation [jyɵ] in several words in the Heshan dialect stands out from all the Taishan-type dialects. Since it is so different from the other dialects (e.g. [v^uɔ] or [vua]), it might involve some further changes that are outside the scope of this dissertation. The origin of this pronunciation awaits further research.

Lastly, I consdiered in Section 7.3 whether the Taishan-type dialects should be further divided into two subgroups based on what the MDS plot shows (see the red dotted line in Figure 16). I have found no evidence in the data for such division.

The findings in the Taishan-type cluster are:

1. ɣ-deletion is also the prevalent sound change for these words; [v-] (Type 2a) is found across the Taishan-type dialects.
2. The reflex [p^h-] found in 'to row' seems to be a local pre-*Qieyun* archaism for the Siyi dialects and therefore it is not considered as an exception for ɣ-deletion.
3. Irregularities of ɣ-deletion (reflex [f-]) have been identified in the Siyi dialects, but not in the Huaxian and Yingde dialects.
4. Huaxian and Yingde dialects are structurally more similar to the Guangzhou-type dialects; they are classified as outliers of the Guangzhou-type dialects.
5. There is no need for an internal division within the Taishan-type dialects.

7.4.3 Maoming-type dialects (brown circle)

In the Maoming dialect, the reflex [ʋ-] (Type 2b) is almost found everywhere. [f-] is found in every instance of the pre-u# context; it is phonologically conditioned and therefore is not an exception to ɣ-deletion. The possible exceptions are shown in Table 16.

Character	Gloss	Pronunciation
桓	pillar	fʊn
混	to mix	kuɐn

Table 16. Possible exceptions to ɣ-deletion in the Maoming dialect

'Pillar' is the only word in the Maoming dialect which shows the reflex [f-] in Table 16. In Figure 19, I have shown the initial reflexes of 'pillar' on a map. There are a handful of reflexes such as [v-], [h-] and [p^h-] scattered mainly near the Pearl River Delta, but these reflexes are the minorities. [w-] and [f-] are the dominating reflexes found in Guangdong for 'pillar'. The [w-] reflex (light blue), a result of ɣ-deletion, is mostly distributed in the north and the [f-] reflex (orange), a result of ɣ-devoicing, is distributed in the southern dialects. Coincidentally, the distribution of [fu] vs. [wu] pronunciations (< MC *ɣo/*ɣu) also shows a north-south

division. This geolinguistic pattern suggests that there might be a link between ɣ-devoicing and geography. The pronunciation of 'pillar' and the presence of pre-u# [f-] in the Maoming dialect show that it is part of the southern dialects in the north-south division.

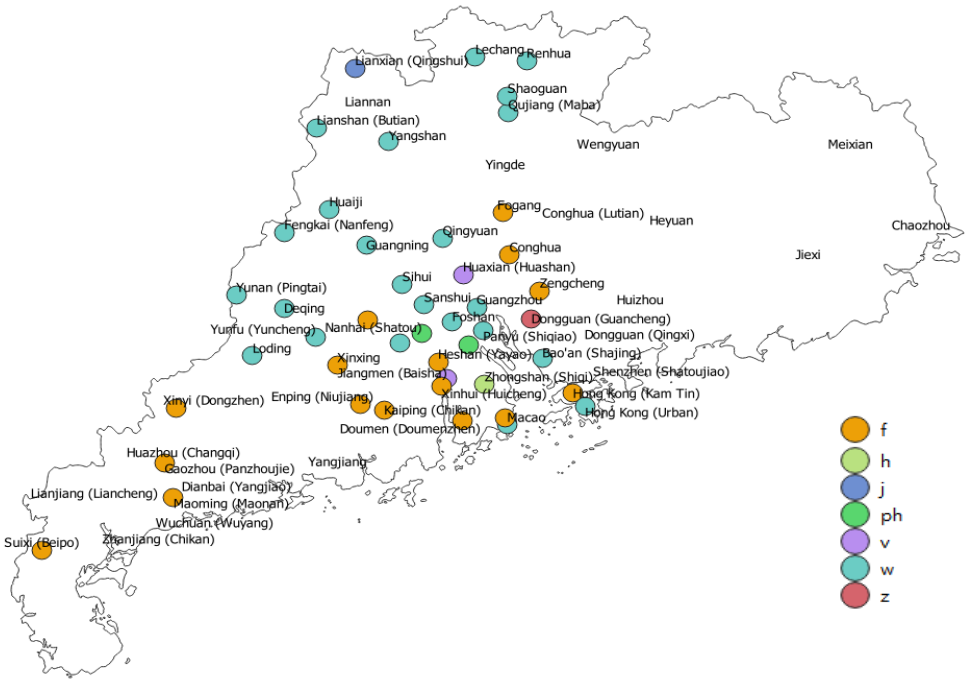

Figure 19. Point-colour map of 'pillar'

Lastly, as Table 16 shows, there is another possible exception to ɣ-deletion in the Maoming dialect. The reflex [k-] found in 'to mix' appears to be a region-specific variant, since it is found in other dialects in the Zhan-Mao area in western Guangdong (light blue circle in Figure 20). I believe that like the [k-] in the central-western Guangzhou-type dialects, this might be pre-*Qieyun* archaism, i.e. having MC *g- (< OC **g-) instead of *ɣ-, for the Zhan-Mao dialects too. Hence, the word does not show the initial [ʊ-] because the MC initial that it corresponds to was not a target of the change in the first place. This item should be excluded in the analysis.

Other Maoming-type dialects include *Dianbai, Huazhou, Wuchuan, Zhanjiang* and *Suixi*. All of these dialects are spoken in the Zhan-Mao area, in western Guangdong. Like the Maoming dialect, the most abundant reflex is [ʊ-], and [f-] is largely phonologically conditioned by _u#.

The possible exceptions are listed in Table 17.

Dialect	Character	Gloss	Pronunciation
Dianbai, Huazhou[2], Wuchuan[2],	混	to mix	kuɐn, kʊɐn[2]
Dianbai,	皖	Anhui	jin
Huazhou, Zhanjiang, Suixi, Dianbai[2]	槐	Pagoda tree	kʰuɐi, kʰʊai[2]
Suixi	桓	pillar	foŋ

Table 17. Possible exceptions to ɣ-deletion in other Maoming-type dialects

The pronunciation of 'Anhui' in Dianbai is addressed in Section 8.1 together with the Guangzhou-type dialects (see Table 12).

The only context that [f-] also appears in is 'pillar', and it was mentioned earlier that this word seems to be linked with the coastal area. However, the form [foŋ] in Suixi is different. This form is different from dialects in the Zhan-Mao area (pronounced as [fun] in other diaelcts). Moreover, in the Suixi dialect, other words are pronounced as [wun] within the same rhyme group, e.g. in 'slow'. 'Pillar' is the only word that stands out with its pronunciation within the rhyme group. Both Mandarin and Hakka in western Guangdong do not pronounce 'pillar' as [foŋ], so it is unclear where this pronunciation comes from.

The reflex [k-] in 'to mix', as mentioned for the Maoming dialect, could be a pre-*Qieyun* archaism and is thus excluded from the analysis. The pronunciations [kʰuɐi]/[kʰʊai] for 'pagoda tree' are also only found in the Zhan-Mao area. I think that like 'to mix', the pronunciation of 'pagoda tree' is an archaism of the region and therefore, it is also excluded in the analysis as well for the same reason as 'to mix'.

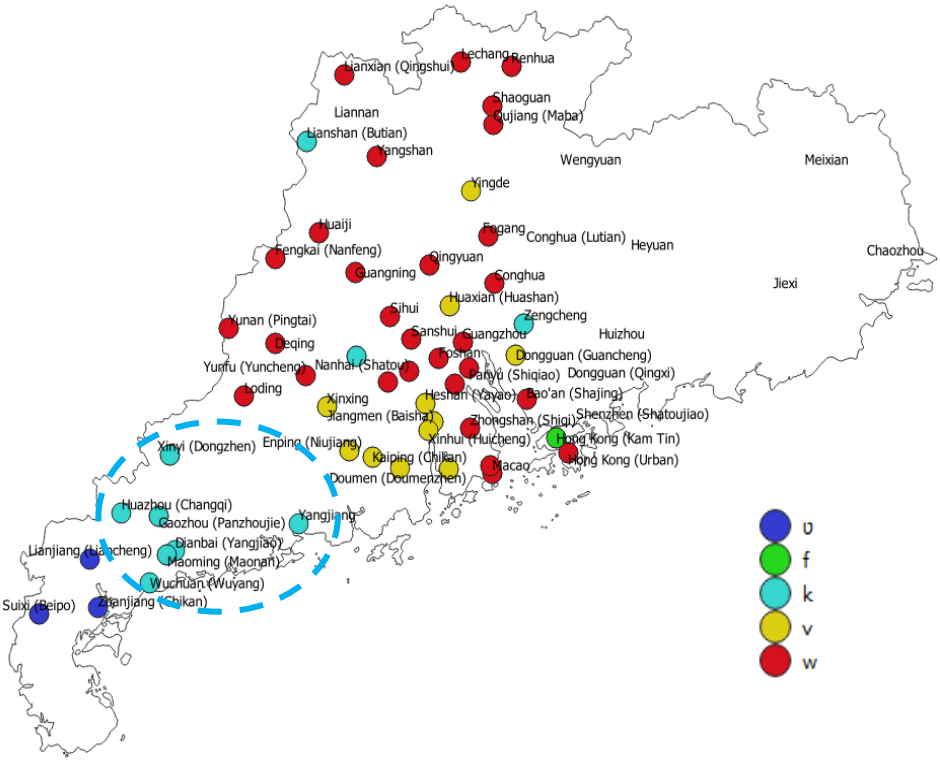

Figure 20. Point-colour map of ‘to mix’

The Wuchuan dialect has a group of words pronounced as [hui]. The [h-] reflex is phonologically conditioned by the rhyme [-ui], therefore it is not an exception to ɣ-deletion. Another interesting observation is that conditioned retention, i.e. the initial [h-] in the [hui] words above, is mainly found in southern Yue dialects too. The Gaoming dialect from the Guangzhou-type dialects shows exactly the same thing, with the presense of [h-] before all the [-uoŋ] rhymes (see Table 12). This phonologically-conditioned rentention supports the idea that there could be a link between ɣ-devoicing and geography. This is addressed in Section 8.2.

Although [ʊ-] is found as the predominant reflex of MC *ɣu-, [w-] can also be found in some contexts. In the Suixi dialect, the reflex [w-] is found phonologically conditioned before the rhyme [-un]. The [wun] pronunciation is not an exception to ɣ-deletion of course, but it shows something interesting. I argue that [-un] is a residual context for the w-develarisation (Type 2b in Figure 8), which I come back to in Section 7.4.5.

The findings in the Maoming-type cluster are:

1. ɣ-deletion is also the prevalent sound change for these words; [ʊ-] (Type 2b) is found across the Maoming-type dialects.
2. [f-] is systematically found in the pre-u# context.
3. Southern Yue dialects tend to have a reflex [f-] for 'pillar', including some Zhan-Mao dialects. Furthermore, the distribution of a [fu] vs. [wu] contrast also shows this north-south division.
4. The [k-] reflex in 'to mix' and [k^h-] for 'pagoda tree' are local pre-*Qieyun* archaisms. They are excluded in the analysis.
5. The Wuchuan dialect shows evidence for systematic ɣ-devoicing before [-ui]. This is very similar to a southern Yue dialect of Gaoming which shows the same retention in another phonological context. I suspect this is linked to the north-south division that I raised earlier.
6. The Suixi dialect retains [w-] from w-develarisation (Type 2b) before a [-un] rhyme.

7.4.4 Bao'an-type diaelcts (blue circle)

The Bao'an dialect can be said to have characteristics of both the Guangzhou-type dialects and the Maoming-type dialects. The most abundant reflex is [w-], but pre-u# [f-] is found systematically. Pre-u# occurrences of [f-] are therefore not considered as exceptions. A possible exception is shown in Table 18 below.

Character	Gloss	Pronunciation
核	core	fɐʔ

Table 18. Possible exceptions to ɣ-deletion in the Bao'an dialect

Similar pronunciations for 'core' are only found in the Guan-Bao dialects in Zhan's (2002) classification; it is pronounced as [fɐt] in Dongguan (Guangcheng) and [fɐk] in Hong Kong (Kam Tin). This can be seen in Figure 17 from Section 7.4.1 (circled in red). Interestingly, in

the Hakka dialects spoken in Dongguan and Shenzhen (near Bao'an, at the border of Hong Kong), 'core' is pronounced as [fut]. Perhaps the [fɐt/-k/-ʔ] pronunciations were influenced by the nearby Hakka dialects. It requires further research to examine their relationships.

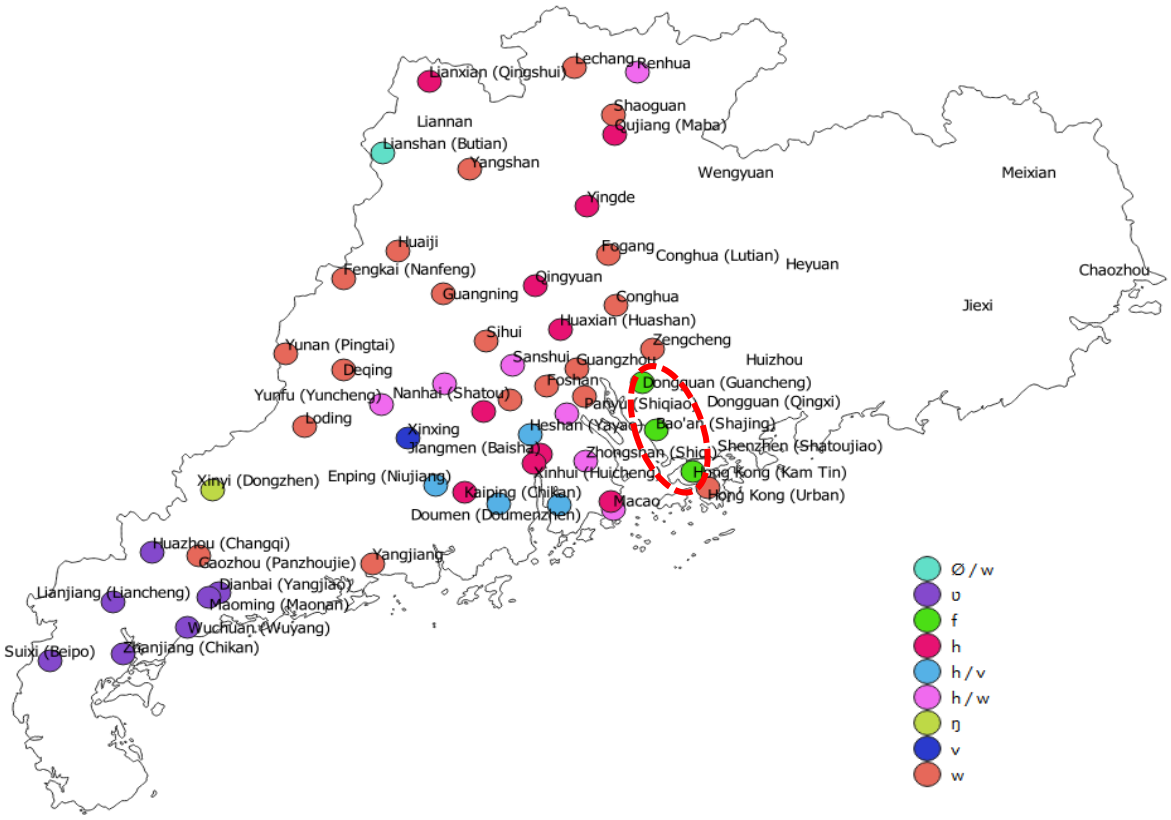

Figure 17. Point-colour map of 'core' (from Section 7.4.1)

Other Bao'an-type dialects include *Hong Kong (Kam Tin), Shunde, Xinyi, Zengcheng, Zhongshan*[41] and *Zhuhai*. All these dialects are spoken near the Pearl River Delta, except for Xinyi, which is spoken in the Zhan-Mao area in western Guangdong. Table 19 below shows the possible exceptions in these dialects. Like the Bao'an dialect, pre-u# [f-] will not be listed.

[41] See Section 7.4.1 for why the Zhongshan dialect belongs to this cluster.

Dialects	Character	Gloss	Pronunciation
Xinyi, HK (Kam Tin)[2], Shunde[3], Zengcheng, Zhuhai	桓	pillar	[fun], [fuŋ][2], [p^{h}un][3],
Shunde, HK (Kam Tin)[2], Zengcheng	緩	slow	[fun], [fuŋ][2],
HK (Kam Tin)	渾	muddy	[fɐŋ]
Xinyi, HK (Kam Tin)[2], Zengcheng[3]	混	to mix	[kuɐn], [fɐŋ][2], [kuɐŋ][3]
Shunde, HK (Kam Tin)[2], Xinyi[3], Zhuhai	核	core	[hɐt], [fɐk][2], [ŋɐt][3],
Shunde	換	change/ switch	[fun]
Shunde	活	to live/ alive	[fut]
Shunde, Zhuhai[2]	餛	‘won’ in ‘Wonton’	[k^{h}uɐn], [k^{h}ɐn][2]
Zengcheng	皖	Anhui	[jun]

Table 19. Possible exceptions to ɣ-deletion in other Bao’an-type dialects

It is interesting to see that Guan-Bao and Xiangshan dialects (Zhan 2002) are close to each other on the MDS plot. It is surprising that the Xinyi dialect, which is located in the Zhan-Mao area, also belongs to this cluster, but the Siyi dialects, which are spoken in the area between Guan-Bao and Zhan-Mao dialects, are distinct from these dialects. There could be a historical relationship between these geographically-distant (Zhan-Mao) dialects and the Bao’an-type dialects. Some evidence can be found in Table 19.

The most outstanding characteristic that can be found in Table 19 is the amount of initial [f-] in these words. The reflex [f-] is already found phonologically conditioned by the pre-u# context. In addition, sporadic [f-] are found in a number of other words; they are not phonologically conditioned in most dialects (Shunde is the only exception of this generalization[42]). In the previous subsections, I have shown that sporadic initial [f-] occurs mainly in southern dialects, often in a similar subset of lexical items in different dialects from

[42] ‘Anhui’ is pronounced as [wun] instead of [fun], but it could borrowed from neighbouring dialects, e.g. Panyu.

different dialect groups. I believe that this is not just a coincidence. Moreover, the systematic pre-u# [f-] can be found in both Bao'an-and Maoming-type dialects. I argue that there is a historical reason for the link between the occurrences of [f-] and geography. A proposal is given in Section 8.2.

The rest of the items in Table 19 appear to be sporadic in relation to the words with the same rhymes. In the Shunde dialect, [p^{h}un] 'pillar' is pronounced identically as the Nanhai dialect. This pronunciation could be a surviving archaism for these two dialects. The pronunciation for 'Anhui' in the Zengcheng dialect is addressed in Section 8.1 with other related forms. For the [k-] reflex in 'to mix', I have already explained that I see them as local archaisms. It seems like 'won' in 'wonton' is a similar case. For [ŋɐt] 'core' in the Xinyi dialect, unfortunately this analysis cannot provide an explanation for why they appear as they are. I make an assumption that it is also a local pre-*Qieyun* archaism. These words are all excluded in this analysis.

Last but not least, my observation of the Zhongshan dialect shows that the pre-u# conditioned reflex is Type 4 ([h-]), and the elsewhere reflex is [w-]. Having an [h-] for this group of words is the main difference between Zhongshan and the rest of the Bao'an-type dialects. Zhan (2002: 201) pointed out that the reflexes of *ɣu- in Zhongshan very much pattern with the Bao'an and Dongguan dialects. The reason why Zhongshan is different in this aspect is perhaps because it was influenced by Hakka and Min dialects, according to Zhan. Although the surface realisation has been changed (to [h-] < *f-) due to dialect contact, the phonological conditioning of the reflex is still shared in these dialects, reflecting their historical links.

The findings in the Bao'an-type cluster are:

1. ɣ-deletion is also the prevalent sound change for these words; [w-] (Type 1) is found across the Bao'an-type dialects.

2. [f-] is systematically found in the pre-u# context. However, there are also numerous words with [f-] where it is not phonologically conditioned.
3. Some of the members in this cluster are geographically very distant. This suggests a historical link between them.
4. Bao'an-type dialects provide more examples for the observation that the reflex [f-] is more commonly found in the south of Guangdong.
5. There are more items, e.g. 'pillar' and 'to mix', which seem to be pre-*Qieyun* archaisms.
6. The Zhongshan dialect shows structural similarity with the rest of the Bao'an-type dialects. The main difference is having pre-u# [h-] instead of [f-], which could be from Hakka/Min influence.

7.4.5 Outliers/ transitional dialects

The remaining dialects that have not been discussed are the outliers in the MDS plot that are not close enough to any cluster to be considered as a member. These dialects include: *Fengkai, Lianshan, Zhaoqing, Nanhai, Gaozhou, Lianjiang, Dongguan* and *Xinxing*. This subsection is an attempt to examine the properties of these dialects in order to evaluate whether they should belong to an existing cluster or they are different enough to be considered as their own groups.

Figure 21. Fengkai and Lianshan dialects on the MDS plot

The Fengkai and Lianshan dialects are close enough to be evaluated together. As shown in Figure 21, the dialects are located near the Guangzhou cluster, but are clearly not part of the same cluster on the MDS plot. The main reason for this seems to be the presence of zero-initials in a number of words, in the Lianshan dialect in particular. These words turn out to be the words which have the [hoŋ] pronunciation in other dialects, such as 'great', 'grand', as well as some

general lexical exceptions like ‘red’, ‘flood’ which are listed in Table 11. The pronunciation of ‘mutual’ in the Lianshan dialect is [n̩], which stands out from the whole Guangdong province. Furthermore, ‘Anhui’ and ‘to mix’ are pronounced as [k^{h}un] and [k^{h}uɐn] respectively. All these differences are enough to make both dialects distant themselves from the Guangzhou cluster on the MDS plot. Based on the pronunciation and the reflexes of the rest of the words, the Fengkai and Liansan dialects resemble the Guangzhou-type dialects. Furthermore, I believe pronunciaitons like [n̩], [k^{h}un] and [k^{h}uɐn] are archaisms.

In the Lianshan dialect, [oŋ] (<*ɣuŋ, e.g. ‘great’ and ‘grand’) is found in a number of words which are usually pronounced as [hoŋ] in other dialects. It looks like ɣ-deletion applied in this particular dialect only, but it turns out ɣ-deletion has nothing to do with it. Through looking at historical minimal pairs with a MC *x-, ‘baked’ 烘 and ‘persuade’ 哄, I found that these syllables also lost their initials. It seems like *ɣuŋ and *xuŋ first merged as *huŋ from ɣ-devoicing and debuccalisation,[43] then the *h- was lost later resulting the [oŋ] form.

I consider both dialects as outliers of the Guangzhou-type dialects.

Figure 22. Zhaoqing and Nanhai dialects on the MDS plot

Figure 22 shows that Zhaoqing is located between the Guangzhou and Bao’an clusters on the MDS plot whereas Nanhai is located below the Bao’an cluster. The Nanhai dialect is like the

[43] An alternative analysis is that there was no debuccalisation; *x- was lost before debuccalisation applied to the rest of the *x- in the dialect. Lowering of the vowel is also omitted here.

Bao'an-type dialects, in that the reflexes of MC *ɣu- are mainly [w-] and [f-], where [f-] is phonologically conditioned by the pre-u# context. In addition, *ɣu- is also reflected as an [f-] before [-un] and [-ut] as well as [-ɔ], [-ɔk] and [-ɔŋ]. This means that, for the Nanhai dialect, the sources of [f-] were more than just the pre-u# position. This explains why Nanhai is located close to the Bao'an cluster, but not part of it. In addition, [f-] is found in parallel distribution with [w-] before [-ak]. This is related to the changes occurred in the southern Yue dialect area which I come back to in Section 8.2. As a whole, I would argue that this dialect is a peripheral member of the Bao'an cluster because it possesses the cluster's main characteristic.

In the Zhaoqing dialect, the reflexes of MC *ɣu- are largely [w-]; [f-] is also found in the pre-u# environment. Unlike the Bao'an-type dialects, the reflex [f-] is not found across all pre-u# contexts. Out of eleven words, only 'household', 'Shanghai', 'mutual' and 'to protect' are pronounced as [wu]; otherwise [fu]. In other words, the [wu] and [fu] pronunciations are in parallel distribution. This pattern resembles a *mixed lect* (Chambers & Trudgill 1998: 110). A mixed lect mixes forms "in an area intermediate between one uniformly has [variant A] and one which has [variant B]" (Chambers & Trudgill 1998: 110). Variants A and B would be [wu] and [fu] respectively here. Zhaoqing is geographically located between the Guangzhou-type dialects (to its north) and the Bao'an-type dialects (Nanhai, in the southeast). Perhaps geographical proximity caused Zhaoqing to have a transitional status between the Guangzhou-cluster and the Bao'an cluster, which the MDS plot (Figure 22) reflects.

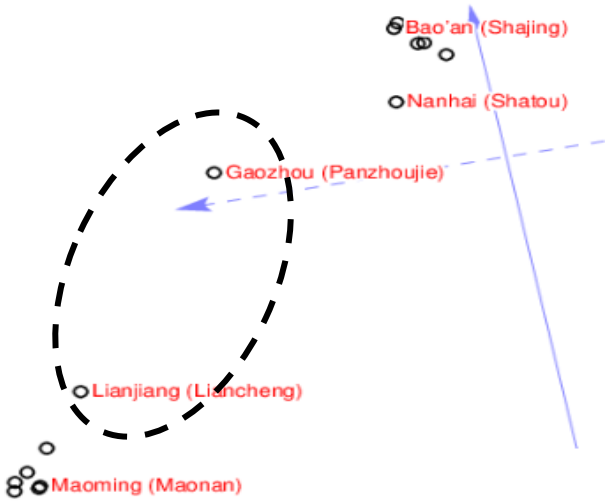

Figure 23. Gaozhou and Lianjiang dialects on the MDS plot

The Gaozhou dialect shows a mixed pattern between the Bao'an-type dialects and the Maoming-type dialects, as reflected in Figure 23. Like both dialect clusters, the Gaozhou dialect also has the reflex [f-] conditioned by the pre-u# context. However, in the elsewhere environment, both [ʋ-] and [w-] are found, sometimes in parallel distribution, e.g. [ʋan] 'fantasy/magical' vs. [wan] 'to suffer from'. Geographically speaking, Gaozhou is situated between a Bao'an-type dialect (Xinyi, with [w-]) and other Maoming-type dialects (with [ʋ-]). The Gaozhou dialect appears to be a mixed lect too, which shows a transitional pattern between the two dialect clusters.

One should also ask, Xinyi is geographically very far away from Bao'an or Zhongshan (Section 7.4.4), so why is the Xinyi dialect closer to the geographically-farther dialects rather than its neighbouring Zhan-Mao dialects? One possibility could be that historically, they are linked. What causes the difference between the Bao'an- and Maoming-type dialects is actually an innovation started in western Guangdong – develarisation of w (see Type 2b in Figure 8). The evidence comes from the Lianjiang dialect data.

The Lianjiang data in Shao (2016) shows that [w-] is conditioned before [-ui], [-un] and [-ut] ('pillar' is the only exception), otherwise, [ʊ-] is found in the rest of the environments. However, this may not have been the case around 15 years ago. Zhan's (2002) data shows that w-develarisation has not started back then.

Zhan (2002) provides additional data that he collected as a supplement to the areas that previous three Yue surveys excluded. One of the locations covered is Lianjiang. The data from the two surveys are compared in Table 20. The [w-] initial is found in Zhan (2002). By the time Shao (2016) did the fieldwork, w-develarisation has already taken place.

Lexicon	Lianjiang (2002)	Lianjiang (2016)
Grain	w̲ɔ	ʊ̲o
Chinese	w̲a	ʊ̲a
Birch	w̲a	ʊ̲a
Lake	fu	fu
Household	fu	fu

Table 20. Additional data of Lianjiang dialects from Zhan (2002)

Zhan (2002) and Shao's (2016) data suggest two things: 1) between 2002 and 2016, a new change, w-develarisation, has occurred in the Zhan-Mao area in western Guangdong and 2) historically the Bao'an-type dialect around the Pearl River Delta and the Maoming-type dialects might be more linked together than we have previously thought. Before w-develarisation, Bao'an- and Maoming-type dialects share the same pattern (i.e. having Type 1 and Type 3 changes, the lattern conditioned by word-final [u]), see Section 8.2. It is likely that they would belong to the same cluster on the MDS plot, even though these dialects are geographically separated by the Siyi dialects. The transitional pattern we see in both Gaozhou and Lianjiang dialects suggests change in progress. This change has completed in most dialects that Shao (2016) surveyed, with the Xinyi dialect being the least affected from this change.

Figure 24. Dongguan (Guancheng) and Xinxing dialects on the MDS plot

Lastly, the Dongguan (Guancheng) and Xinxing dialects are situated near the Taishan cluster (see Figure 24). The Dongguan dialect is the only dialect that possesses [v-] and [f-] as major reflexes of *ɣu-, where [f-] is conditioned by word-final [-u]. The conditioned [f-] is a characteristic of the Bao'an- and Maoming-type dialects. w-fortition (Type 2a) is a recent independent innovation[44] in the Dongguan dialect which made it much closer to the Taishan cluster instead of the geographically-closer Bao'an cluster on the MDS plot. The Dongguan dialect is like the Yingde and Huaxian dialects; it shows surface similarities (abundant reflex [v-]), but structurally, it is closer to another dialect clusters. I consider it as an outlier of the Bao'an-type dialects.

The Xinxing dialect is much more similar to other dialects in the Taishan cluster. Reflex [v-] is pretty much found in all contexts, with no pre-u# [f-]. What could have made this dialect stand out in the MDS plot is the reflex [f-] in 'red', 'vast/flood', 'large' and 'rainbow', which is different from the [h-] in the Taishan-type diaelcts.[45] Other than these four words, the Xinxing dialect basically patterns with the Taishan cluster. Therefore, it belongs to the Taishan-type dialects based on its structural similarity.

The findings in the outliers on the MDS plot are:

1. Most dialects belong to their nearby cluster on the MDS plot (as outliers), except Zhaoqing, Gaozhou and Dongguan dialects.
2. A recent independent w-fortition in the Dongguan dialect caused high surface similarity

[44] In Wang & Qian's (1949) study, the Dongguan dialect did not have [v-] in its inventory, only [w-].
[45] See Table 11 and Section 8.1.4.

(shown as small distances on the MDS plot) with other Taishan-type dialects. The pre-u# [f-] shows its structural similarity with the Bao'an-type dialects, which makes sense, geographically speaking.

3. Zhaoqing and Gaozhou dialects are mixed lects. There are different reflexes found in parallel distribution in the same phonological context, which show their transitional character. The geographical proximity explains their transitional status.
4. The Nanhai dialect shows a much higher instances of conditioned [f-] than its neighbouring dialects. However, it is still considered as a Bao'an-type dialect since it has the main characteristics of the dialect cluster.

In Table 10, I have proposed a provisional typology of reflexes of MC *ɣu-. The MDS plot has proved to be a great assistance in finding dialect patterns in the analysis. A revised typology of the dialect patterns is shown in Table 21. There are four types of dialect patterns, as suggested by the MDS plot, with outlier dialects (e.g. Dongguan) and transitional dialects (e.g. Gaozhou).

Representative dialect	**Type**	**Traditional dialect classification (*LAC* 2012: 127)**
Guangzhou	Type 1	Guangfu, Goulou
Taishan	Type 2a	Siyi
Bao'an	Type 1 & Type 3 (Phon. Conditioned)	Guangfu
Dongguan (Guancheng) -under Bao'an-type	Type 2a & Type 3 (Phon. Conditioned)	Guangfu
Zhongshan -under Bao'an-type	Type 1 & Type 4 (Phon. Conditioned)	Guangfu
Maoming	Type 2b & Type 3 (Phon. Conditioned)	Gao-Yang
Zhaoqing -Transitional dialect	Type 1 & Type 3	Guangfu
Gaozhao -Transitional dialect	Type 1, 2b & Type 3 (Phon. Conditioned)	Gao-Yang

Table 21. A revised typology of dialect patterns according to their reflexes of MC *ɣu-

8. Discussion

In Section 7.4, I have found that: 1) all dialects had ɣ-deletion across various phonological environments, 2) some dialects show restriction of ɣ-deletion in certain environments, e.g. pre-u#, which led to ɣ-devoicing, 3) in some dialects, pre-*Qieyun* archaisms can be found and 4) reflex [f-] is largely found in southern Yue dialects. Section 8.1 assesses the general lexical exceptions from Table 11 and considers whether they play any roles in the irregularity of ɣ-deletion. Section 8.2 gives an analysis on the origins of [f-], what the distribution of this reflex can tell us about ɣ-deletion and the formation of the dialect landscape regarding to this feature.

8.1 General lexical exceptions

I mentioned at the start of Section 7.4 that I have taken out 14 words (see Table 11) from the data to be analysed separately because they repeatedly appear to be very different from the rest of the data across different dialect groups. Table 11 is given again below. Multiple reflexes indicate geographical (sometimes intra-dialectal) variation for the relevant word.

Reflex	Word	Gloss	Reflex	Word	Gloss
f	乎	Interrogative particle	h/f	宏	grand
f	晃	to sway	h/k	紅	red
j/z	完	to finish	h/k	洪	flood
j/z	丸	meat ball	h/k	鴻	large
k^h/k	潰	to rot	h/k	虹	rainbow
k^h	繪	to paint	h/k/k^h	汞	mercury
h/f	弘	great	h/k	鬨	coax

Table 11. General lexical exceptions

The following subsections discuss these words in more detail in order to determine whether they are exceptions to ɣ-deletion or not. I start with the left hand side of the table.

8.1.1 'Interrogative particle' and 'to sway'

For these two items, we would expect to see [wu] and [wɔŋ] in the Guangzhou dialect, for example. However, they are found reflected with an [f-] initial repeatedly in almost all dialects.

The interrogative particle is mainly found in Classical Chinese texts. It could have entered Yue as a literary pronunciation, from a variety that did not have ɣ-deletion. 'To sway' is also not a native word in Yue. [jiu (pai)] is usually the word in colloquial Cantonese. It seems likely that 'to sway' is another borrowing which only exist as a literary pronunciation. These words fit Lau's (2003) theory[46] that [f-]-initial-words are borrowed.

8.1.2 'To finish' and 'meat ball'

There are two main types of reflexes for these two words, [z-] for some Siyi dialects and [j-] for most other dialects.

The reflex [j-] occurs before front high vowels, [y] or [i], as in [jyn] and [jin] (and [jun] in the Zengcheng dialect) for 'to finish' and 'meat ball'. It is important to know whether the initial [j-] is a reflex of MC *ɣ- or not. Chen & Newman (1984a: 188) proposed that ɣ-procope occurred before MC medials *-u-, *-i- and *-y-, then followed by glide-insertion. If we hypothesise that the earlier form of [jyn] and [jin] are *yn and *in, they would fit the environments for ɣ-procope. If this is true, then, [j-] is simply the result of glide-insertion (Chen & Newman 1984a: 188) after ɣ-procope. In this hypothesis, [j-] is not a reflex of *ɣ-.

If [j-] is not a reflex of *ɣ-, what about [z-] in the Siyi dialects? The pronunciation for 'finish' and 'meat ball' is [zin] in the Doumen dialect, but it is [zᵘɔn] in the Taishan dialect. Zhan (2002: 143, 146) shows that there is a systematic correspondence between Guangzhou [j-] and Siyi [z-] as well as Guangzhou [w-] and Siyi [v-]. Based on this description, we would expect to see [z-] occurring before *-i- and *-y- and [v-] before *-u- in the Siyi dialects. Initial [z-] in [zin] in the Doumen dialect seems to be a fortition of [j-] in [jin], which fits Zhan's description. This is not the case for the Taishan dialect, though. [z-], which we expect to see preceding [i] or [y], is followed by a [ᵘ].

[46] See Section 5.2.3.

Early records of the Taishan dialect show that [v-] and [z-] are recent innovations. The records of the Llin-nen dialect in the late 19th century (Don 1882, 1883) and the Taishan dialect in 1950 (Wang & Qian 1950b) show no signs of [v-] and [z-]; only [w-] and [j-][47] were recorded. Based on these descriptions, *j^{u}ɔn is reconstructed for these words in Early Siyi dialects. The next question is regarding to the correspondence between Early Siyi *j- and MC *ɣ-. Is *j- in Early Siyi a reflex of MC *ɣ-? To address this question, I have hypothesized two possible correspondences of these segments, visualized in (17a) and (17b) below:

(17a)

MC	ɣ	u	ɑ	n
Early Siyi	j	^{u}	ɔ	n

(17b)

MC	ɣ	i	u	ɑ	n
Early Siyi		j	^{u}	ɔ	n

(17a) hypothesizes that that there is a change directly from *ɣ- > j-. This hypothesis only states a phonological change that could have happened, it does not provide any better explanation as to why and how it happened. It is also unclear why the change only affected these two words. Moreover, phonetically speaking, I am skeptical to the change *ɣ- > j-. It is unlikely that a dorsal consonant preceding a back vocoid would become a palatal segment. (17a) seems very implausible.

An alternative hypothesis is shown in (17b). The glide *-i- was there all along. It looks very odd to have two glides being the medial of a word in the Grade I category in a rhyme group. If we look at the historical minimal pairs in the IV category within same rhyme group (i.e. *-u- vs. *-iu-), we can see that in most dialects they are identical (in the *Shan* rhyme group), see

[47] Wang & Qian (1950b: 72) noted that [j-] is sometimes pronounced as [ʒ-] or [ʒj-]

(18). I argue that 'finish' and 'meat ball' are examples of words which do not fit the description of Middle Chinese rhyme books such as *Guangyun*. Instead, they should be treated as Grade IV words in the *Shan* rhyme group[48], supported by their present-day pronunciation.

(18)

Dialect	Gloss	Pronunciation	MC Vowel grade
Guangzhou	to finish	jyn	I
	mysterious	jyn	IV
Taishan	to finish	$z^{u}ɔn$	I
	mysterious	$z^{u}ɔn$	IV

If we do not follow the description of these words in *Guangyun*, the origins of these two irregular forms are immediately clear. Using his dialect data, Chen (2011: 151) has proposed the following changes in the rhyme from MC to the present-day Guangzhou dialect pronunciation, shown in (19a). Based on Chen's reconstruction, I have reconstructed the pathway of change for the Taishan dialect. The Taishan dialects took a different, slightly more conservative route (regarding to the medials), see (19b).[49] Two significant differences between the Guangzhou and Taishan changes are: i) no coalescence of medials *-i- and *-u- and ii) the nucleus vowel was backed in Taishan. The changes in (19b) resulted in the drastically different pronunciations we see in (18), where the medial *-u- is retained. This explains why [z-] (< *j-), expected to precede *-i-, is found before [^{u}] here (which applies to 'meat ball', 'to finish' and other Grade IV words). This account also explains why the Zengcheng dialect has the form [jun] instead of [wun] for these words.

(19a) *iuæn > yœn > yøn > yn
(19b) (*iuæn? >) *iuɑn > *$j^{u}ɔn$ > $z^{u}ɔn$

[48] See Table 5 in Section 3.2.4 for the reconstruction.
[49] I am not certain whether the initial starting point of the change in Taishan would be identical to Guangzhou (*iuæn), so I put a question mark after it.

Lastly, this hypothesis also predicts that for words that did not have medial *-i- would get a present-day [v-], which is what we would have expected. This is indeed the case. In the word ‘to change/switch’ for example, the pronunciation is [v^{u}ɔn] (< *ɣuɑn). After ɣ-deletion, only the medial *-u- was left, followed by glide-insertion of *w-. Until recently, when w-fortition led to [v-]. The presence of [j-] is therefore not irregular.

The discussion in this subsection includes words that share the same form with ‘finish’ and ‘meat ball’ (i.e. [jVn]), e.g. ‘Anhui’ and ‘slow’ in some Guangzhou-type and Maoming-type dialects.

8.1.3 ‘To rot’ and ‘to paint’

The following discussion focuses on tonal reflexes and the data is drawn from dialects in the Pearl River Delta region only. To recap: each MC tonal category has a tone reflex (in tonal contours). Moreover, Yue tones were split into *Yin* and *Yang* categories after MC obstruent devoicing. The split was based on the phonation of the MC initial (voiceless > *Yin*; voiced > *Yang*). See Section 3.2.5 for more.

‘To rot’ and ‘to paint’ had *ɣ- with the *Qu* tone, according to *Guangyun*. Therefore, we expect to see the tone reflex for *Yang Qu* for this group of words. The actual tonal reflex of these two words suggests something else, however. In many dialects, both words are found in the *Yin* category, meaning that when the tone split happened, the initial was voiceless. This means that unlike the previous archaism explanation, these initials are not the result of MC obstruent devoicing since the *Yin* category does not support their earlier form to be *g-. What is even more interesting is that the tone contours that these words bear reflect the *Shang* category in some dialects, not the *Qu* category. Some pronunciations of ‘to rot’ are shown in Table 22. It is

assumed that this can be extended to account for ‘to paint’ too since in many dialects these two word are homophones.

If *Guangyun* truly reflected the earlier stage of Yue dialects, the tones for the words listed in Table 22 should all reflect the MC *Qu* category. This is not the case. I argue that the phonological history of ‘to rot’ and ‘to paint’ is not as is assumed in the traditional description. The early forms of these words in the Yue dialects were different from the MC variety recorded in *Guangyun*, which is not surprising.

Dialect	Pronunciation	Tone	MC Tone category correspondences	Register
Guangzhou	k^hui	35	*Shang*	*Yin*
Hong Kong	k^hui	35	*Shang*	*Yin*
Panyu	k^hui	35	*Shang*	*Yin*
Huaxian	k^hoi	35	*Shang*	*Yin*
Zengcheng	k^hɔi	35	*Shang*	*Yin*
Foshan	k^hui	35	*Shang*	*Yin*
Nanhai	k^hu	13	*Shang*	*Yang*
Shunde	k^hui	32	*Qu*	*Yin*
Sanshui	k^hui	33	*Qu*	*Yin*
Doumen	k^hui	55	*Shang*	*Yin*
Dongguan	k^hui	35	*Shang*	*Yin*
Bao'an	k^hui	24	*Shang*	*Yin*

Table 22. Tonal reflexes and MC tone category correspondences for ‘to rot’

Let us assume for now that the proto-initial of ‘to rot’ is *x-, the voiceless counterpart of *ɣ-, since the tone reflects the *Yin* category. I have found two words which are homophones of ‘to rot’ in Middle Chinese (again from *Guangyun*), i.e. words pronounced as *xuɑi (in the *Shang* category). These words are ‘to bribe’ *賄* and ‘to regret’ *悔*. We would expect the present-day pronunciation of these words to be something like [fui] (from *xu- > f-) and for them to have

a tonal reflex of the *Shang* category. 'To regret' is pronounced with [f-] as expected, but 'to bribe' is pronounced [k^h-] in the Guangzhou, Hong Kong, Panyu, Nanhai dialects. What is more interesting is that the tone for 'to regret' (with [f-]) actually corresponds to MC *Qu* category and not *Shang* (i.e. again different from *Guangyun*). This means that for these four words, if the word corresponds to the *Shang* category, we get [k^hui], otherwise, [fui]. This is the case found in 'to bribe', 'to rot' and 'to paint'. This pattern suggests that in some dialects, *x- experienced fortition and became [k^h-] in a certain phonological context, illustrated in (21a). Alternatively, [k^h-] could be the earlier form and spirantisation occurred before the same segmental context with the *Qu* tone category instead (see 21b). Further research is needed to find out the exact direction of change for these items and to what extent to they apply to all dialects.

(21a) *x- > k^h- / _ui#, *Shang* tone
(21b) *k^h - > x- / _ui#, *Qu* tone

This is just an initial observation of the tone-sound change correspondences. The discussion above is to show that these two words should not be considered as exceptions to ɣ-deletion because there is very good reason to believe that they show an alternative phonological history to the one that the rhyme book has recorded. It also shows an interesting change which is outside the scope of this dissertation and is worth exploring further in the future. Although this finding only applies to some dialects in the PRD region, it can certainly shed some light on other areas in Guangdong, where [k^h-] is almost found everywhere.

8.1.4 'Great' and 'grand'

'Great' and 'grand' are often homophones in each dialect group. Furthermore, they show regional variation for their reflexes too. [w-] is the most common reflex (from ɣ-deletion), found in Guangdong. Reflexes [f-] and [h-] are the other two major initial reflexes for these words.

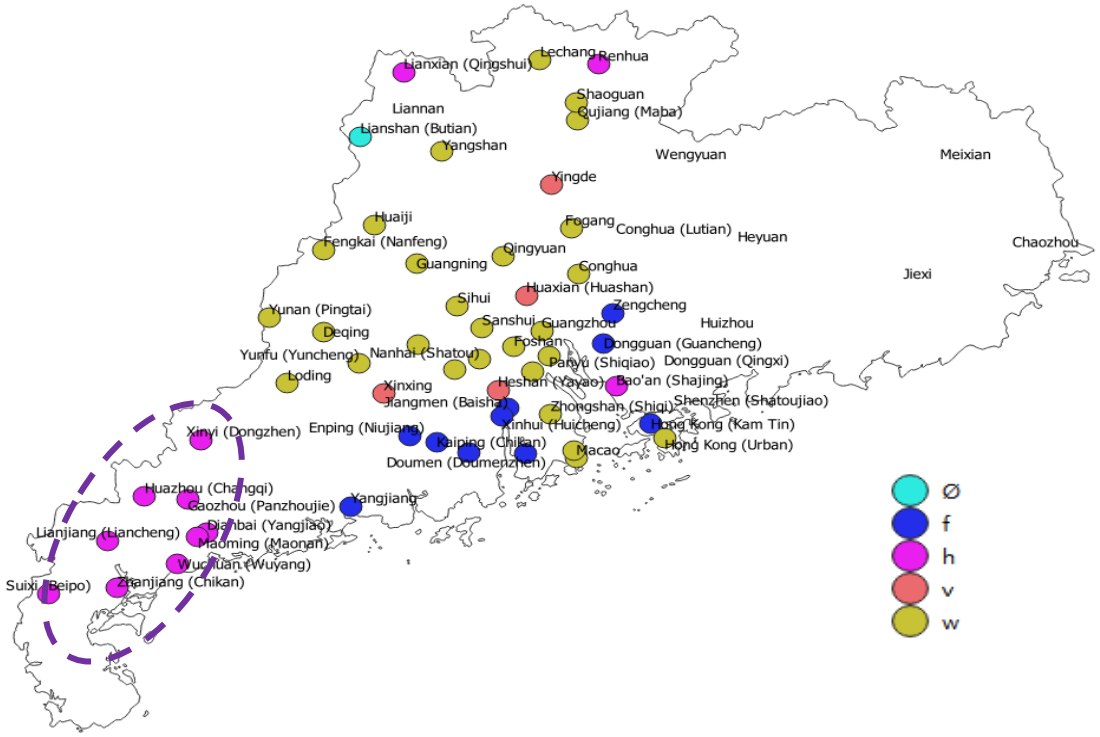

Figure 25. Point-colour map of ‘grand’

Figure 25 shows that reflex [w-] is mainly found in central and northern Guangdong, as well as in some Pearl River Delta dialects. These words are often pronounced as [wɐŋ]/[waŋ] in these dialects. The [h-] reflex is found in some Zhan-Mao dialects, the Bao’an dialect and two dialects in northern Guangdong (circled in purple in Figure 25); these words are often pronounced as [hoŋ]. The geolinguistic distribution of the reflexes suggests that the [hoŋ] pronunciation could be a result of a separate development from [wɐŋ]/[waŋ]. The [hoŋ] pronunciation could originate from Guangxi. The Guangxi dialect data (Chen & Lin 2009: 323, 351) shows that the dialects on the other side of the border in the west also exhibit the [hoŋ](/[huŋ]) pronunciation. It is necessarily to examine the Guangxi data to find the historical changes for these two items.

Lastly, reflex [f-] is found in the Siyi dialects for these words. This addressed in Section 8.2 together with other instances of sporadic [f-].

8.1.5 The [hoŋ] words

In Table 11, the remaining words are pronounced [hoŋ] or [koŋ] in most dialects. Based on the present-day pronunciation, these words do not show traces of MC medial *-u-, even though they are marked as *closed rhymes*. There are two possible explanations for their pronunciations: 1) borrowing and 2) involvement of other sound changes.

Figure 26 shows the geographical distribution of the initial reflexes of 'red'. It shows that almost all dialects have the initial [h-] for this word. This is also the basic pattern for 'flood', 'large', 'rainbow' and 'coax'. I would argue that with such a low number of alternative reflexes, it is unlikely that the [h-] pronunciation is a borrowing. For the [k-] initial, I believe that it is like reflex [k-] 'fox', 'to mix' from before, that it could be a local archaism.

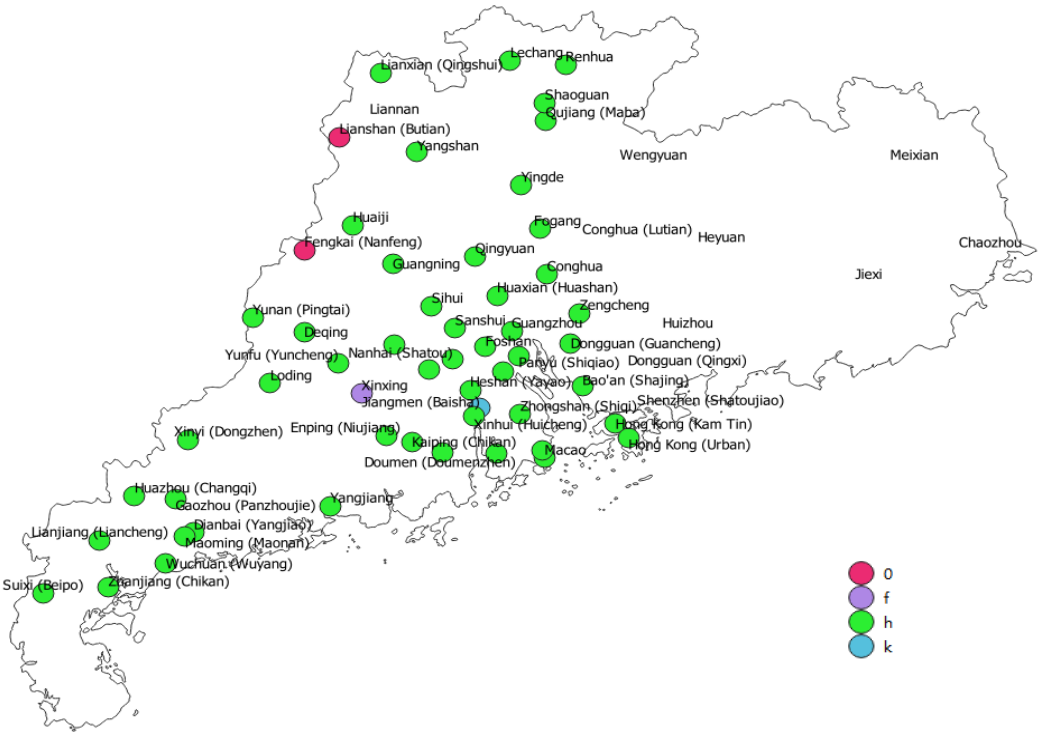

Figure 26. Point-colour map for 'red'

Figure 27 shows the distribution of reflexes of 'mercury', which shows much more variation than Figure 26. The geolinguistic pattern is difficult to spot, as there are no clear region-specific

variants. Major variants, [h-] and [k-], are found in every dialect cluster on the MDS plot. However, the [h-] reflex tends to be more common at central Guangdong and the PRD area. 'Mercury' is a likely candidate for borrowing. However, it is unclear which reflex is borrowed based on the current data.

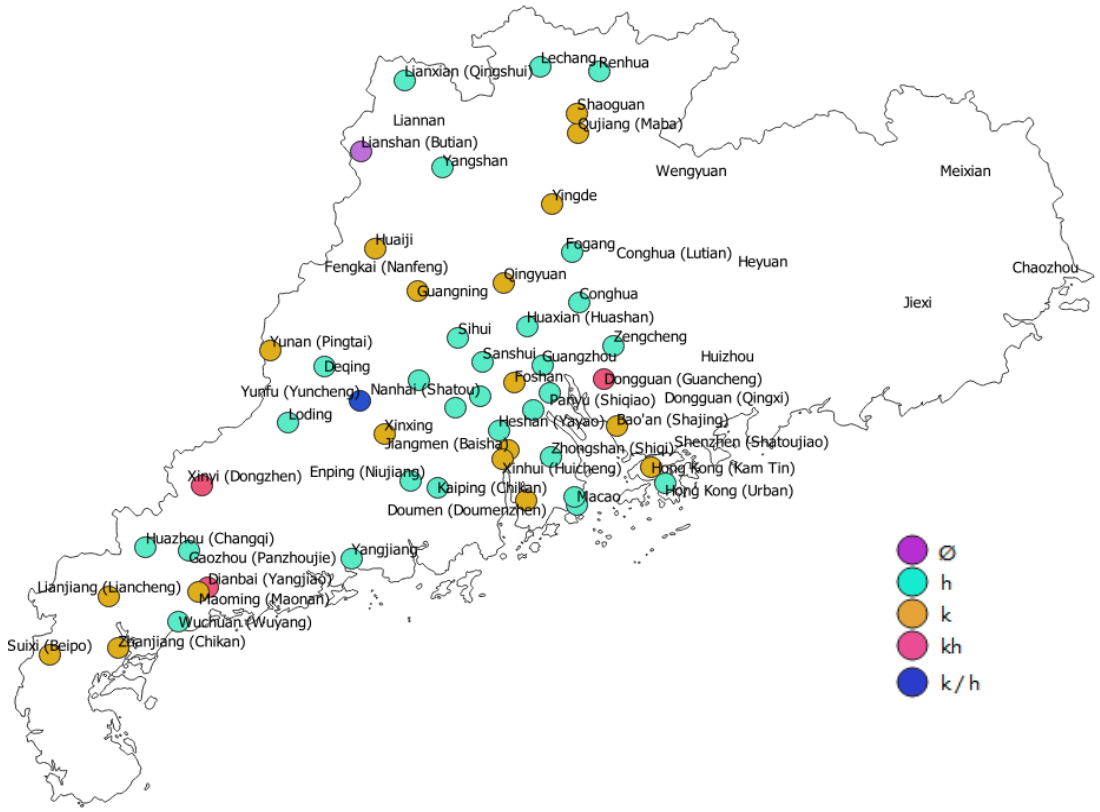

Figure 27. Point-colour map for 'mercury'

Present-day [-oŋ] in these words corresponds to the MC *Tong* rhyme group, which is reconstructed as *-uŋ. Since rhyme books only show the rhyme categories and not their sound values, if *-uŋ only changed its sound value but it remained distinct from other rhyme groups, it would stay as a separate rhyme group. This phonetic change therefore would be unrecorded.

If we accept the *-uŋ reconstruction, then what happened between MC and now should have involved u-lowering before *-ŋ across the dialects. Considering the relative chronology of this change, it probably occurred before ɣ-deletion. In other words, the lowering of *u bled ɣ-

deletion, which produced the [hoŋ] form as a result of ɣ-devoicing. This is visualized in Table 25 in Section 8.2.2. In this case, words pronounced as [hoŋ] should be separated from the analysis of ɣ-deletion, since when the change happened, the phonological context in these words was already irrelevant.

8.1.6 Putting things together

All the lexical items assessed above should be excluded from the analysis of ɣ-deletion. This is because these items are either borrowings (e.g. ‘to sway), words that did not meet the requirements to participate in the sound change (e.g. ‘red’) or words that are outside the scope of this dissertation (e.g. ‘to end’).

It is now appropriate to say that ɣ-deletion is essentially regular in most Yue dialects. The [f-] initial is mostly responsible making ɣ-deletion seem irregular in some dialects. Here I am referring to the sporadic words with the [f-] initial, which is most abundant in the Taishan-, and Bao’an-type dialects. Now that I have pointed out there are sporadic cases of [f-] in these dialect groups, one must ask why that is the case. I argue that Taishan-type dialects used to be like Bao’an-type dialects, but it was swamped by a newer stratum which is related to the Guangzhou-type dialects. This hypothesis explains why: 1) other than the [v-] vs. [w-], structurally Taishan- and Guangzhou-type dialects are very similar, 2) the Siyi dialects has broken the Maoming- and Bao’an-type dialects, which seem to share a lot of linguistic features as they could be geographically linked together, and 3) sporadic cases of [f-] are also found in Siyi dialects. I provide more evidence for this hypothesis in Section 8.2. This hypothesis predicts that the sporadic cases of [f-] are in fact archaisms too.

Lastly, in Section 8.1, I have proposed that i) [f-] comes from borrowings of a historical standard variety; ii) [k^h-] comes from a sound change, and iii) [j-] is part another group of

words which this dissertation is not concerned with. However, the former two explanations go against what I have established in Section 7.4, that 1) [f-] comes from ɣ-devoicing, which has something to do with geography and 2) [k-] is possibly pre-*Qieyun* archaisms.

The way I distinguish what I wrote in Section 7.4 and Section 8.1 is based on the geographical distribution of these reflexes. The conclusions I made in Section 8.1 is based on the fact that these reflexes are found in almost all Yue dialects in Guangdong. The chances for the survival of archaisms in almost all dialects is very small. Moreover, [k^h-] is predictable in a number of dialects by the phonological change I stated in Section 8.1.3. Perhaps this is not a coincidence at all. On the other hand, reflexes such as [k-] and [p^h-] (Section 7.4) are usually found in isolation or in a specific region. There is also not enough evidence to support that they were innovations. Archaism is the most sensible explanation for them.

For [f-], 'to sway' and 'interrogative particle' (from Section 8.1) are items found in classical texts. I have reasons to believe that they are borrowings. Contrastingly, it would be very strange to find a north-south division for the [fu] pronunciation as a borrowing, especially when it is phonologically systematic. I argue that it is more sensible to postulate that the geolinguistic distribution of [f-] for the words in Section 7.4 reflects an older dialect area and the origin of [fu] is a systematic sound change. These are elaborated in Seciton 8.2.

8.2 The origins of reflex [f-] and the formation of present-day dialect landscape of Yue in Guangdong

The reflex [f-] is present at different rates in different groups of Yue dialects spoken in Guangdong. In some dialects, [f-] is found conditioned by the pre-u# context (e.g. Maoming); some dialects show that [f-] can be conditioned by more than one environment (e.g. Nanhai); in other dialects, [f-] is found sporadically (e.g. Taishan). Perhaps coincidentally, [f-] is found in dialects spoken in Western Guangdong and around the Pearl River Delta, next to the Hakka-

speaking region, where it is common to find [f-] reflex for MC *ɣu-. In this section, I first examine the Hakka-contact hypothesis in the data, then I give an alternative explanation for distribution of reflex [f-] in the data.

8.2.1 Contact with Hakka dialects

Gan & Shao (2000: 50) suggest that reflex [f-] was a result of the contact between Yue and Hakka. The authors claim that a number of Yue dialects show similarities with Hakka with the evidence of ɣ-devoicing instead of ɣ-deletion, unlike the Guangzhou dialect. This hypothesis aimed at the dialects spoken in the Pearl River Delta (Siyi dialects, Guan-Bao dialects as well as Huaxian, Conghua, Zengcheng, Nanhai, Shunde, Sanshui and Zhongshan dialects), but it could in theory be applied to other areas of Guangdong, where Hakka dialects are also spoken, e.g. the Zhan-Mao area. Therefore, this hypothesis can potentially explain why dialects from the Zhan-Mao area as well as the Guan-Bao dialects have [f-].

According to the *SDPRD* (Zhan & Cheung 1990: 22), the Hakka dialects in the Pearl River Delta area (Dongguan (Qingxi), Shenzhen (Shatoujiao), Zhongshan (Nanlang), Conghua (Lütian) and Huizhou) have both [f-] ([h-] in Huizhou) and [v-] as reflexes of MC *ɣu-, although the phonological environments they occur in are not specified. Delving deeper into the Hakka data shows that some dialects are similar to each other (like the Dongguan and Shenzhen dialects), but not all of them are as homogenous. The Dongguan and Shenzhen dialects show numerous instances of [f-], before various contexts including [-u], [-a], [-ui] and [-ai]. There are other positions where [f-] is also found, but they are in parallel distribution with [v-]. Table 23 gives some examples of words in which [f-] is found in three Hakka dialects spoken in different parts of Guangdong.

Dialect	Gloss	Pronunciation
Dongguan	Chinese	fa
	Lake	fu
	Return	fui
	Bad	fai
	To live/ alive	fat
Zhongshan	Chinese	fa/ va
	Lake	fu
	Return	fui
	Bad	fai
	To live/ alive	fat
Xinyi	Lake	fu
	Return	fui
	Bad	fai

Table 23. Examples from several Hakka dialects spoken in Guangdong

Table 23 does not show the full picture of for the Hakka dialects, as the examples given above barely scratch the surface. There seems to be more variation in the contexts that ɣ-deletion and ɣ-devoicing could have occurred in the Hakka dialects than in Yue; the exact historical changes of MC *ɣu- in Hakka is another study of its own. In general, ɣ-devoicing in Hakka is not as phonologically conditioned as the Yue dialects. Therefore, we cannot be 100% sure which language, Yue or Hakka, influence the other language for now.

Since both the southern Yue dialects near the Pearl River Delta and the Hakka dialects there have the reflex [f-], Gan & Shao's (2000) contact-hypothesis is not entirely impossible. However, if that is the case, what drove the borrowings from Hakka to only affect a certain class of words, i.e. words with a pre-u# context, but not the others? Thus, there are no exceptions in pre-u# words having an initial [f-] in many dialects. This seems too regular for borrowings. What is even more interesting is that I have found a non-pre-u# word ('slow') which in Yue has a reflex [f-], but in Hakka dialects it has a reflex [v-]. The example is shown in Table 24 for Dongguan Yue and Hakka:

Gloss	Dongguan Yue	Dongguan Hakka
slow	[fun]	[vɔn]

Table 24. ‘Slow’ in Dongguan Yue and Hakka

According to Gan & Shao’s hypothesis, we would expect Dongguan Hakka to influence Dongguan Yue with its [f-] in the word ‘slow’. However, there is no [f-] in Dongguan Hakka for ‘slow’. The directionality of influence does not add up. There is no immediate evidence that the pronunciation [fun] in Dongguan Yue is an influence from Hakka.

Some of the dialects which have a higher proportion of the reflex [f-], namely the Siyi dialects and the Nanhai dialect, are not spoken in/next to an area where Hakka would be expected to cause an influence. According to *Language Atlas of China* (Wurm et al. 1989: B15), Hakka dialects are “thinly scattered” in the Siyi (and a large portion of the Guangfu) area. It is unclear how much influence Hakka could actually have on the urban Yue dialects in these areas. The most likely cases which Hakka could have an influence to Yue dialects are the sporadic [f-] found in Qujiang, Fogang and Renhua, since they are Yue islands surrounded by Hakka varieties. It is more likely to see Hakka influence in this kind of sociolinguistic situation than the rest of the Yue dialects discussed.

Based on the lack of explicit evidence from Gan & Shao and the lack of explanation for the systematicity of the pre-u# [f-] reflexes, their hypothesis is not very convincing in most cases. I propose an alternative hypothesis for the origins of the reflex [f-] in Section 8.2.2.

8.2.2 The history of the rhyme [-u]

In Chen & Newman’s (1984a: 188) analysis, they mention that MC *u did not trigger ɣ-procope, but *o did. By the time ɣ-procope was active, MC *u was already diphthongized and *o had

raised to a [-u] (see Section 5.2.3). I argue that this is the key to understanding why [f-] is conditioned by the pre-u# position in certain dialects, but not in others.

Firstly, as I brought up many times in Section 7.4, the reflex [f-] is mainly found in the south, near the coast. The phonological conditioning of [f-] is also only found in the south, though this generalization excludes the Siyi dialects. The geographical correlation with the phonological conditioning may not be a coincidence after all. (22) is a simulation of what might have happened for both the northern Yue dialects and the southern dialects. (22) is based on the assumption that present day [-u] comes from MC *o (c.f. Chen & Newman 1984a).

(22)

	Stage 1	Stage 2	Stage 3	Stage 4
Northern Yue	*ɣo	*ɣu	wu	wu
Southern Yue	*ɣo	*xo	*hu	fu

In (22), both dialect areas started with *ɣo (Stage 1). In the northern Yue dialects, o-raising occurred (Stage 2) and triggered ɣ-deletion, presumably before ɣ-devoicing reached these locations. Glide-insertion was applied following ɣ-deletion (Stage 3) and this is the form we see in the Guangzhou dialect (Chen & Newman 1984a: 188). The northern Yue dialects basically followed Guangzhou's changes.

On the other hand, I assume that o-raising had not reached the southern Yue dialects; ɣ-devoicing instead of ɣ-deletion was triggered (Stage 2) because in these words, the phonological condition was not right for ɣ-deletion then. After these words experienced ɣ-devoicing, o-raising finally spread to the south and raised all the occurrences of *o, then *x was debuccalised (Stage 3). Lastly, labio-dentalisation affected all the *hu sequences, which thus became [fu] (Stage 4).

The hypothesis above predicts two things: a) o-raising came from the north and b) o-raising is a neogrammarian change. o-raising affected the northern dialects first and did not reach the south in time for ɣ-deletion to apply in these words. This is the evidence for o-raising being a northern innovation (in Guangdong). o-raising also affected all the words with the same vowel all at once, without any exceptions (following *ɣ-). This suggests that o-raising was clearly a neogrammarian change.

Piecing all the information together, Table 25 summarises where and why ɣ-deletion applied, in a relative chronological order, based on what I discussed in the previous sections. The [fu] pronunciation in the south is therefore a Yue innovation, not an influence from Hakka (as Gan & Shao 2000 suggest).

Dialects	Northern Yue	Southern Yue	Both	Both
Rhyme group	Yu (*o)	Yu (*o)	Other rhyme groups	Tong (*uŋ)
Stage 1	*ɣo	*ɣo	*ɣu + rhyme	*ɣuŋ
Stage 2	*ɣu	*ɣo	*ɣu + rhyme	*ɣoŋ
Stage 3	wu	*xo	w + rhyme	*xoŋ
Stage 4	wu	*hu	w + rhyme	hoŋ
Stage 5	wu	fu	w + rhyme	hoŋ

Table 25. Relative chronology of the development of MC *ɣu- in different rhyme groups

8.2.3 Remnant forms

What I showed in Section 8.2.2 still does not explain the sporadic [f-] found in the southern dialects. The Macau dialect is an excellent starting point for tackling this problem. The Macau dialect is very similar to the Guangzhou dialect, except that it has more sporadic reflex [f-]. This is linked to the formation of the Macau dialect.

In the late 1870s, the variety spoken in Macau was closely related to present-day Zhuhai and Zhongshan dialects, which are spoken near Macau (Lin 1998: 502, cited from Lo 2013: 154).

By 1941, the Macau dialect already showed a structural resemblance with the Guangzhou dialect. Lo found that the Macau dialect has shifted from a Zhongshan-type dialect to a Guangzhou-type dialect in just the 44 years since 1897 (Lo 2013: 64). Lo believes that this structural change was not caused internally, but by contact. Luo points out that the population and demographics changed drastically at the beginning of the 20th century. Lo believes that the structural changes occurred around the 1920s. The population in 1920 was 84,000, but by 1927, it rose to 157,000; the population kept rising until after WWII (Lo 2013: 155-156). This change in demographics unsurprisingly changed the language use in Macao. The Guangzhou dialect became a 'koine' for dialect speakers who moved to Macau (Lo 2013: 158). This new variety of Macao Yue eventually replaced the old variety and it was stabilized after the mid-20th century (Lo 2013: 156). The Guangzhou dialect acted like a new 'stratum' (Mai 2009, see Section 4.3.2), which swamped the the old Macau dialect, leaving only a small amount of the old dialect layer, i.e. words with the [f-] reflex. This is the explanation for the occurrences of sporadic [f-] in the dialect and its similarity with the Guangzhou dialect.

Before considering another dialect, I propose two names to refer to the dialects with different features. I call the dialects with pre-u# [f-] *Coastal* Yue dialects, since they are spoken closer to the coast in southern Guangdong than the other group of dialects. Note that the Coastal dialect group does not include the Siyi dialects, since they do not share the same feature. I call the dialects spoken in northern and central Guangdong the *Inland* Yue dialects, in contrast to the Coastal dialects. The Inland dialects include Siyi dialects.

Following the analysis of Macau, having a few sporadic words with the [f-] initials might shed some light on patterns in other dialects, as well as the historical dialect landscape. In the Sanshui dialect, there are two words with an initial [f-] as a reflex of *ɣu- (excluding 'to sway'). 'Grain' is pronounced as [fɔ]. This pronunciation is only found in Sanshui and Nanhai. The

other word is ‘birch’ [fa]. This time, the same pronunciation is found in the Siyi dialects (Jiangmen and Heshan). While one could argue that ‘grain’ was borrowed from Nanhai because they are neighbouring dialects, the same statement cannot be applied to ‘birch’, because the Siyi dialects are not contiguous to Sanshui. Both the Macau and Sanshui dialects have sporadic instances of reflex [f-]. Perhaps like Macau, the Sanshui dialect used to have more reflexes with [f-], like ‘grain’ and ‘birch’, and it was swamped by a new stratum too. Sanshui is located near the border of the Inland and Coastal dialects. The Sanshui dialect could possibly belong to the Coastal dialects historically. In the past several hundred years, the prestige of Guangzhou might have expanded outward and driven the Sanshui dialect to converge to the Inland dialects. This is a possibility, since the further away one gets from the Inland dialect area, the more [f-] reflexes can be found in the PRD region, like the Bao’an-type dialects. It suggests the effect of the Guangzhou dialect reduces as we go further away.

In the southern Guangzhou-, Bao’an-, Taishan- (and to a lesser extent Maoming-) type dialects, there are more varieties like Sanshui which show different degrees of the [f-] reflex other than those which are preceding [-u]. If we hypothesize that these are in fact remnant forms from the earlier Coastal dialects, then we could estimate the historical extent of the early Coastal dialect. The estimated historical area is discussed in the next subsection. Lastly, the term ‘remnant’ here suggests that the sporadic appearances of the reflex [f-] as the result of contact: the newer stratum swamped the older dialect, leaving a handful of archaisms behind. This is explored in Section 8.2.5.

8.2.4 The Inland vs. Coastal dialects

I shall now define more clearly what the Inland vs. Coastal dialects are. Based on the dialect survey data and the assumption that the sporadic [f-] reflexes are remnant forms (see Section 8.2.3), the Coastal dialects refer to dialects which mostly 1) have the reflex [f-] in pre-u#

contexts (see Section 8.2.2) and 2) possess (sporadic) [f-]-initial word forms (sometimes [h-]) other than 'to sway' and the interrogative particle. The extent of the present-day Coastal dialects goes from the coastline of Guangdong all the way up to Zengcheng, Shunde, Nanhai and Xinyi, excluding the Siyi area. The Inland dialects on the other hand are dialects that largely lack the features listed for the Coastal dialects. This is illustrated in Figure 28 below.

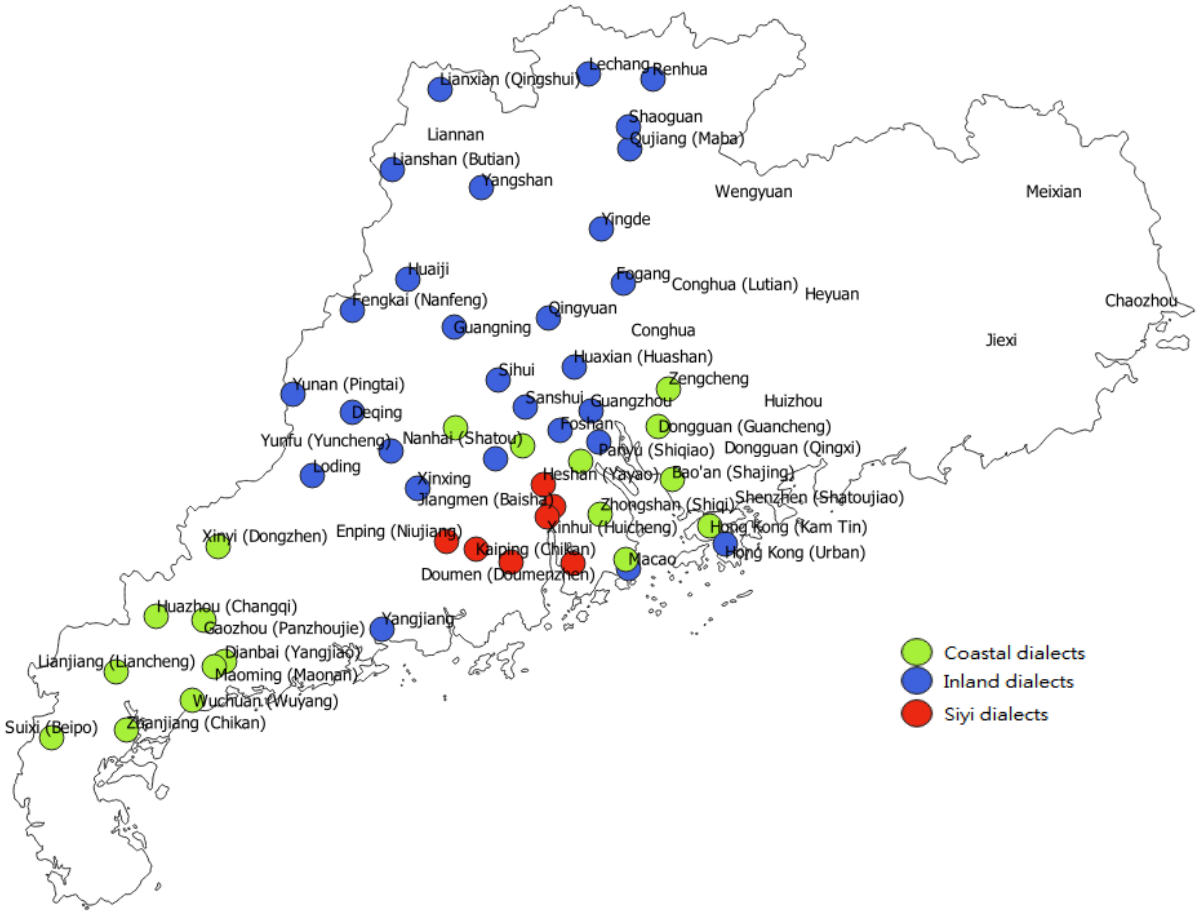

Figure 28. Map of Inland, Coastal Yue and Siyi dialects in Guangdong

Figure 28 shows the Inland vs. Coastal dialect division I have been discussion in Section 8.2. The Coastal dialects are indicated in light green. On the other hand, the Inland dialects are in blue. The north-south division is clearly illustrated in Figure 28. Moreover, I have purposefully separated the Siyi dialects (in red) in order to show they do not belong to the Coastal dialects. Hong Kong and Macau are marked as Inland dialects because they are transplanted varieties in the Coastal area. I have previously mentioned that Yangjiang shares some sporadic items (e.g.

some instances of [fu]) like the Zhan-Mao dialects. I suspect the reason for this dialect to resemble the Inland dialects is similar to the Siyi dialects, which I explain below.

Figure 28 shows clearly that the Siyi dialect region and Yangjiang break the Coastal dialect area into two blocks: the Pearl River Delta (PRD) block and the Zhan-Mao block. The Zhan-Mao dialects show strong ties with the PRD block since both groups have a shared innovation of late o-raising (see (22)), and have the pronunciation of 'pillar' as [fun]. The discontinuous distribution of these features in southern Guangdong raises questions on how this geolinguistic landscape was formed. There are two possibilities, migration from the Pearl River Delta to the Zhan-Mao area or a historical dialect area being interrupted. I argue for the latter below.

The Siyi dialects basically pattern with Guangzhou, except Guangzhou [w-] corresponds to [v-] in the Siyi dialects (see Section 7.4.2 for the description). However, there is also a surprising number of similarities between the PRD Coastal dialects and the Siyi dialects. These are listed in (23)-(25) below.

(23)

	Enping	Kaiping	Taishan	Jiangmen	Xinhui	Doumen
'great'	faŋ	faŋ	faŋ	fɐŋ	faŋ	faŋ
'grand'	faŋ	faŋ	faŋ	faŋ	faŋ	faŋ

(24)

	Zengcheng	Dongguan	Hong Kong (Kam Tin)
'great'	hɐŋ	fɐn	fɐŋ
'grand'	fɐŋ	fɐn	fɐŋ

(25)

	Enping	Kaiping	Taishan	Jiangmen	Xinhui	Doumen
'household'	fu	fu	fu	fu	fu	fu
'Shanghai'	fu	fu	fu	fu	fu	fu

(23) shows the pronunciation of 'great' and 'grand' in several Siyi dialects and (24) shows the pronunciation of the same words in three PRD Coastal dialects. Both groups share the basic word form, which is only found in these regions in the whole Guangdong province. This suggests that, historically, the Coastal Yue dialects possibly: 1) were spoken in the Siyi area too and 2) [f-] was a reflex in more than just one context, possibly all. In addition, a few pre-u# words are also pronounced as [fu] in the Siyi dialects, see (25). This further supports the Coastal dialect hypothesis. A possible explanation of these [f-] forms is that they are remnants from swamping of the Inland dialects.

In the Zhan-Mao dialects, 'pillar' [fun] is found outside the pre-u# context having the reflex [f-]. I believe that it is also a remnant.[50] Based on the evidence from Siyi dialects and the Zhan-Mao dialects, I am proposing a hypothesis that the whole southern coastal region of Guangdong was dominated by the Coastal Yue dialects at a certain point in history. The main feature of the Coastal dialect is that they experienced ɣ-devoicing across all phonological contexts instead of ɣ-deletion, unlike the Inland dialects.

8.2.5 Formation of the present-day dialect landscape

Based on the phonological conditions, instances where ɣ-devoicing is found in the dialect data and the present-day geolinguistic patterns, the changes in the dialect landscape configuration regarding to ɣ-deletion and ɣ-devoicing are reconstructed in Figures 29-31 below.

[50] See Sections 7.4.4 and 7.4.5 also for discussion of the historical relationship between the Zhan-Mao and other southern dialects

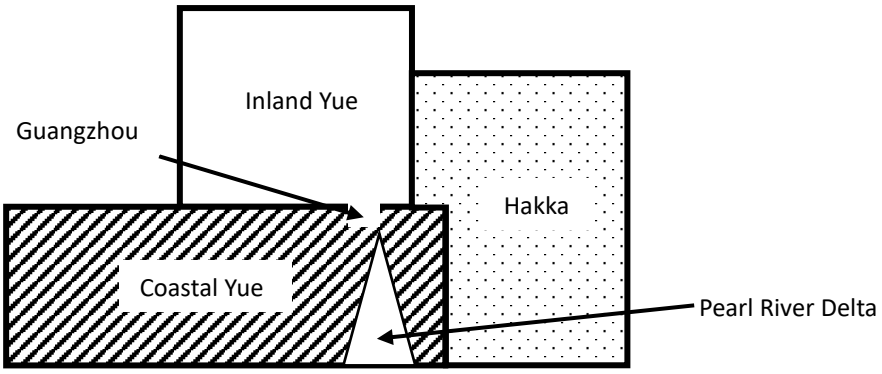

Figure 29. Stage 1. Original dialect configuration

Figure 29 is an abstract map of Guangdong. It shows the hypothesized early dialect areas. Inland Yue represents dialects in the north where ɣ-deletion was applied in all environments. In the Coastal Yue dialects, ɣ-devoicing occurred in pre-u# contexts due to the late arrival of o-raising as well as in all other environments. During this period, labio-dentalisation has not occurred; ɣ-devoicing has produced the *xu-/*hu- sequence. Hakka dialects were spoken in eastern Guangdong. Guangzhou has no shadings in Figure 29 to show that it patterns with Inland Yue.

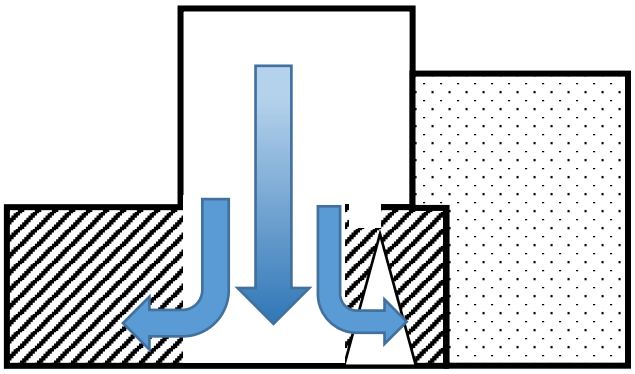

Figure 30. Stage 2. Expansion of the Inland dialects

Figure 30 shows the expansion of the Inland dialects into the Coastal dialect area. The Inland expansion is the source of the new stratum, which caused swamping in Coastal dialects. The result of this is the split of the eastern (Bao'an-type dialects) and western (Zhan-Mao dialects) Coastal dialect blocks; each block has weakened features retained from the early Coastal dialects. Labio-dentalisation is assumed to have occurred in all Yue dialects around this time.

The most salient and unaffected feature that the Coastal dialects have retained from ɣ-devoicing is the reflex [f-] in the pre-u# context (after o-raising). The Siyi and Yangjiang dialects are the most affected dialects in the whole region. Even the characteristic words with pre-u# [f-] were almost all wiped out, leaving only a handful of remnant forms behind and a handful of words in other environments with reflex [f-]. The arrows in Figure 30 are illustrations of the influence from the Inland dialects. The exact origin and the route of the wave still requires further research.

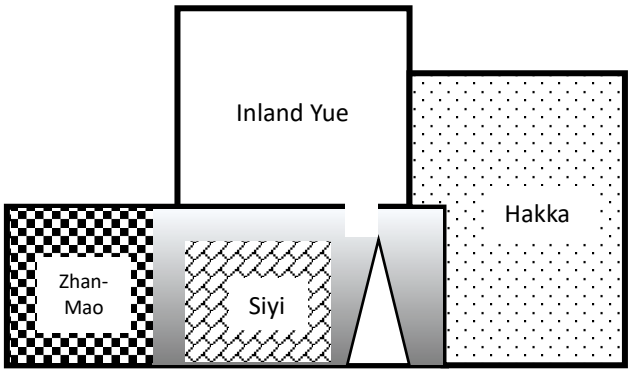

Figure 31. Stage 3. Later changes and the dialect landscape in present time

As indicated by the new shading in the Coastal dialect area, Figure 31 shows how the older stratum/features were 'eroded' and these dialects became more similar to the Inland dialects. The gradual change of colour in the shading represent the degree of 'erosion' as the distance from the Inland dialects increases. The darker the shading, the more retention it has. An example of this 'erosion' is the Sanshui dialect, which preserves only two words with the Coastal reflex [f-], whereas the Bao'an dialect further south keeps the reflex [f-] in pre-u# words as well as numerous other words. Lastly, the Siyi and Zhan-Mao dialects each have different shadings. This is to show that they developed their new reflexes of *ɣu- differently. The Siyi dialects had w-fortition (Type 2a) and the Zhan-Mao dialects had (and some are still implementing) w-develarisation (Type 2b). The lighter shading of the Siyi dialects also indicates the loss of pre-u# [f-] compare to other Coastal dialects.

My proposal for the origins of the reflex [f-] is different from previous studies in a few aspects: 1) occurrences of [f-] are actually remnants from the old Coastal dialect group which were swamped by the Inland dialects, 2) occurrences of [f-] are not exceptions to ɣ-deletion as Li (1996) suggested and 3) in the Coastal dialects, occurrences of [f-] were neither an influence from Hakka (as Gan & Shao 2000 suggest) nor a borrowing from Guanhua as Lau (2003) proposed; ɣ-devoicing was native to the Coastal dialects. Future studies should consider exploring whether there are other Yue features which support the Inland vs. Coastal dialect areas and the swamping effect.

There are some final questions regarding the formation of the present-day dialect landscape. When did ɣ-deletion occur? Lau (2003) believes that the zero reflex [Ø-] (through ɣ-deletion) reflects the phonology before the Tang dynasty. Does this mean the Coastal dialects were imported varieties after Tang? If so, when did the importation happen and which historical variety did it split from? Next, assuming that ɣ-devoicing affected the whole historical Coastal dialect group, when did the expansion of the Inland Yue dialects happen? Would this be related to the migration wave in the late Song dynasty (Lau 2001) or some other demographic changes that did not catch the eyes of historical phonologists since they were not influential enough? Or, perhaps it was more of a diaglossia situation instead of migration, which eventually led to dialect attrition? Next, the Zhan-Mao dialects show a lesser extent of retention of the Coastal dialect features; is this because the data for them were collected much later than for other dialects, and the prestigious Guangzhou dialect had already asserted its influence in the area and these other varieties were already almost levelled? Lastly, why did the pre-u# [f-] survive the Inland dialect expansion? Was local/regional identity involved? All these questions cannot be addressed in this dissertation unfortunately. It will require future research to consider the model properly and to discover if there is more evidence from demography and/or history for the model proposed in this dissertation.

9. Conclusion

In this dissertation, I have pointed out that there is no systematic, cross-dialect account for the history of the MC *ɣu- sequence in Yue dialects and the variation that these dialects currently show for this feature in the previous literature. My aim in this dissertation has been to find out what kind of inter- and intra-dialectal variation exists in the reflexes of MC *ɣu- in the Yue dialects in Guangdong. To do so, I have used current dialectological methods to analyse exisitng dialect survey data. My approach has gone beyond the traditional analysis of Chinese dialects, which mainly describes the correspondences between Middle Chinese sound categories and present-day dialect reflexes.

Through a close analysis of all relevant lexical items in 54 dialects, I was able to answer the research questions that I raised at the start of this dissertation. Previously, there is no geolinguistic analysis for the dialect variation of the reflexes of *ɣu- in the literature. The two dialect-specific analyses of Guangdong and Dongguan provide a good basis for the analysis of this Middle Chinese sequence. The application of dialectometry in Section 7.3 has proven to be successful at identifying geolinguistic structures and it shows that there are four main dialect patterns. This answers my first research question, which concerns with the existence of geographical patterns of this feature in Yue dialects. My second research question concerns whether the same reflex in found across all the words. The close analysis in Section 7.4 shows that not all the words share the same reflex. My third research question concerns the number of changes involved from Middle Chinese to present-day varieties. Primarily, dialects show ɣ-deletion, which resulted in a [w-] reflex (due to later glide-insertion/formation). Some dialects had a further innovation in recent decades: the Taishan-type dialects had w-fortition and the Maoming-type dialects had w-develarisation. For dialects that also show ɣ-devoicing, they were followed by labio-dentalisation, resulting reflex [f-]. My fourth research question concerns the phonological conditioning for the sound changes. I found that reflex [f-] is found

systematically in the pre-u# environment in two southern dialect groups. In addition to the main patterns, there are a number of lexical exceptions and pre-*Qieyun* archaisms preserved in numerous dialects. These words did not cause irregularity to ɣ-deletion, though. Moreover, I have found two dialects with transitional status, which show characteristics of a mixed lect. In particular, the Zhaoqing dialect shows patterns from two dialect clusters, which can be explained by its geographical proximity between these two dialect groups.

My final two research questions concern whether contact (e.g. with Hakka dialects) was responsible for the dialect landscape and whether there is a historical explanation for dialects which are geographically distant, but very similar linguistically. To address these research questions, I have considered the reflex [f-] (pre-u# and sporadic) in more detail. Firstly, I found that Gan & Shao's proposal that the reflex [f-] came from Hakka dialects is not supported by a sufficient amount of evidence. I proposed an alternative hypothesis. The occurrences of systematic (pre-u#) and sporadic reflexes as [f-] are mainly distributed in the southern dialects. I believe that this is not a coincidence. o-raising did not reach the southern dialects in time for ɣ-deletion and that resulted ɣ-devoicing instead, leading to what we see now as [fu] in present-day dialects. The sporadic [f-] came from the historical Coastal Yue dialect group. These sporadic occurrences of [f-] are remnants of a process of swamping – a new stratum covering the older dialect. This proposal explains the similarity between geographically distant dialects (Bao'an- and Maoming-type dialects) as well as why sporadic occurrences of [f-] are found in southern dialects (namely Siyi dialects), showing that the reflex [f-] is neither an exception to ɣ-deletion as Li (1996) has proposed nor are most of them borrowings, as Lau (2003) has suggested.

Lastly, this dissertation attempts to account for as much variation as possible, but there are still individual variants that can yet be explained. Furthermore, there are still a lot of unanswered

questions regarding to the formation of the present-day dialect landscape and the phonological history of ɣ-deletion in Guangdong. These questions await for further research to address.

References:

Ball, J. D. (1890). The Tung-Kwun (東莞) Dialect. *The China Review*, 18(4): 284-299.

Bauer, R., & Benedict, P. (1997). Modern Cantonese phonology (Trends in linguistics. Studies and monographs ; 102). Berlin ; New York: Mouton de Gruyter.

Beijing University Research Institute of Language (1989). *Han yu fang yin zi hui 漢語方音字匯 [The Lexicon of Chinese Dialects]*. Beijing: Wen zi gai ge chu ban she.

Bo, W. (1995). Gan yu an yi fang yan xia mu zi du yin de li shi ceng ci ji yi zhu xiang guan ben zi de kau shi 贛語安義方言匣母字讀音的歷史層次及一組相關本字的考釋 [The Pronunciation of Xia mu category of words in the Anyi Gan dialect and an analysis of a group of related words]. *Cahiers De Linguistique Asie Orientale*, 24(2): 229-244.

Chambers, J. K. & Trudgill, P. (1998). *Dialectology (2nd Edition)*. Cambridge: Cambridge University Press.

Chan, A. (2006). On the Principle of the Four Grades. In Branner, D. P. (Ed.). (2006). *The Chinese Rime Tables: Linguistic philosophy and historical-comparative phonology*, 37-46. John Benjamins Publishing.

Chang, G. (2019). *Han yu yu yin fa zhan shi 漢語語音發展史 [A Developmental Phonological History of Chinese]*. Taipei: Taiwan Commercial Press.

Chen, H. & Lin, Y. (2009). *Yue yu ping hua tu hua fang yin zi hui di 1 pian: Guangxi yue yu, Guinan ping hua bu fen 粵語平話土話方音字彙第 1 編: 廣西粵語、桂南平話部分 [The Lexicon of Yue, Pinghua and Tuhua Volume 1: Guangxi Yue and Guinan Pinghua]*. Shanghai: Shanghai Education Publishing.

Chen, M. (1976). From Middle Chinese to Modern Peking. *Journal of Chinese Linguistics*, 4(2): 113-277.

Chen, M. Y., & Newman, J. (1984-1985). From Middle Chinese to Modern Cantonese (Part 1, Part 2 & Part3). In *Journal of Chinese Linguistics*, 12(1): 148-198, 12(2): 334-388, 13(1): 122-170.

Chen, W. (2011). *Guangzhou di qu yue fang yan yu yin yan jiu 廣州地區粵方言語音研究 [A Phonological Study of the Yue dialects in the Guangzhou Area]*. Guangzhou: Jinan University Press.

Cheng, C. (1997). Measuring Relationship among Dialects: DOC and Related Resources. In *Computational Linguistics & Chinese Language Processing*, 2(1): 41-72.

Chinagate. (2009). *Guangdong sheng li shi gai mao 广东省历史概貌 [A brief history of the Guangdong Province]*. [Retrieved Dec, 2019, from http://cn.chinagate.cn/culture/kejia/2009-12/04/content_19009663.htm].

DeLancey, S. (2011). Sino-Tibetan Languages. In Comrie (ed), *The World's Major Languages* (Second Edition), 693-702. Oxon: Routledge.

Don, A. (1882-83). The Llin-nen (新寧) variation of Cantonese. *The China Review*, 11: 236-247, 12: 474-481.

van Driem, G. (2014). Trans-Himalayan. In Owen-Smith, T. & Hill, N. W. (eds) *Trans-Himalayan Linguistics*, 11-40. Berlin: Mouton De Gruyter.

Embleton, S. (1993). Multidimensional scaling as a dialectometrical technique: Outline of a research project. In Köhler, R. & Rieger, B. B. (eds). *Contributions to quantitative linguistics*, 267-276. Dordrecht: Springer.

Gan, Y. & Shao, H. (2000). Shi lun ke jia fang yan dui yue yu yu yin de yin xiang 試論客家方言對粵語語音的影響 [On the influence on Yue dialects from Hakka dialects]. *Jinan xue bao: zhe xue she hui ke xue ban 暨南學報: 哲學社會科學版 [Jinan Journal: Philosophy and Social Sciences Edition]*, 22(5): 49-56.

Goebl, H. (2010b). Dialectometry and quantitative mapping. In Lameli, A., Kehrein, R. & Rabanus, S. (eds) *Language and Space. An International Handbook of Linguistic Variation*, vol. 2, 433-457. Berlin: Mouton De Gruyter.

Handel, Z. (2015), Non-IPA Symbols in IPA Transcriptions in China. In Sybesma, R., Behr, W., Gu, Y., Handel, Z., Huang, C., & Myers, J. (eds), *Encyclopedia of Chinese Language and Linguistics*. [Retrieved in Jun, 2020, from http://dx.doi.org.ezproxy.is.ed.ac.uk/10.1163/2210-7363_ecll_COM_00000292].

Heeringa, W. (2004). Measuring dialect pronunciation using Levenshtein distance. Groningen: University Library Groningen. [PhD dissertation, University of Groningen. Available at https://www.rug.nl/research/portal/en/publications/measuring-dialect-pronunciation-differences-using-levenshtein-distance(6ff6bbca-842f-4a90-9c6e-a0d3cce748da).html]

Hill, N. (2019). The Historical Phonology of Tibetan, Burmese, and Chinese. Cambridge University Press.

Karlgren, B. (1949). *The Chinese Language: An Essay on its Nature and History*. Ronald Press Company. Cited in Zhang, G. (2019). *Han yu yu yin fa zhan shi 漢語語音發展史 [A Developmental Phonological History of Chinese]*. Taipei: Taiwan Commercial Press.

Kloss, H. (1967). 'Abstand Languages' and 'Ausbau Languages'. *Anthropological Linguistics*, 9(7): 29-41. [Retrieved Jan, 2020, from www.jstor.org/stable/30029461].

Labov, W. (2001). *Principles of linguistic change Volume 2: Social factors*. Oxford: Blackwell.

Lau, C. (2001). Xian dai yue yu yuan yu song mo yi min shuo 現代粵語源於宋末移民說 [Modern Yue comes from the migration in the late Song dynasty – a hypothesis]. In Lau, C., *Xianggang ke yue fang yan bi jiao yan jiu 香港客粵方言比較研究 [A Comparative Study of Hakka and Yue in Hong Kong]*, 188-200. Guangzhou: Jinan University Press.

Lau, C. (2003). Yue yu yuk e yu wen bai yi du de bi jiao 粵語與客語文白異讀的比較 [A Comparison of Literary vs. Colloquial Pronunciations in Yue and Hakka]. In Macau Chinese Linguistic Society (Editing body), Di 6 jie yue fang yan yan tao hui lun wen ji 第六屆粵方言研討會論文集 [Proceedings of the 6th Yue Dialect Conference]. Cited in Chen, W. (2011). *Guangzhou di qu yue fang yan yu yin yan jiu 廣州地區粵方言語音研究 [A Phonological Study of the Yue dialects in the Guangzhou Area]*.

Li, C. N. & Thompson, S. A. (2011). Chinese. In Comrie (ed), *The World's Major Languages* (Second Edition), 703-723. Oxon: Routledge.

Li, L. (2015). *Dongguan yue yu yu yin yan jiu 東莞粵語語音研究 [A Phonological Study of the Dongguan Yue]*. Guangzhou: Jinan University Press.

Li, R. (1973). *Qieyun yin xi 切韻音系 [The Phonological System of Qieyun]*. Wending Book Company. Cited in Zhu, J. (2016). *Sheng yun xue 聲韻學 [Phonology (2nd Edition)]*. Taipei: Wunan Publishing Company Ltd.

Li, R. (2007). Han yu fang yan xue (di 2 ban) 漢語方言學 (第二版) [Chinese Dialectology (2nd Edition)]. Higher Education Press.

Li, R. (2017). *Han yu fang yan diao cha 漢語方言調查 [Fieldwork for Chinese Dialects]*. Beijing: Commercial Press.

Li, X. & Xiang, M. (2010). *Han yu fang yan xue ji chu jiao qing 漢語方言學基礎教程 [A Foundational Course in Chinese Dialectology]*. Beijing: Beijing University Press.

Li, X. (1996). Yue yu yu gu yin 粵語與古音 [Yue and Ancient Phones]. In *Xue shu yan jiu 學術研究 [Academic Research]*, 8: 74-78.

Li, X. (1997). Shu bai nian lai yue fang yan yun mu de fa zhan 數百年來粵方言韻母的發展 [The Developments of the Rhymes in Yue in the Past Several Hundred Years]. In Li, X., *Li xin kui yin yun xue lun ji 李新魁音韻學論集 [Papers on Phonology by Li Xinkui]*, 413-428. Swatou: Swatou University Press.

Liang, F. (1980). Zhongguo li dai hu kou, tian di, tian fu tong ji 中國歷代戶口、天地、田賦統計 [Statistics of Historical Chinese Hukou, Tiandi and Tianfu]. Shanghai: Shanghai People's Press. Cited in Lau, C. (2001). Xian dai yue yu yuan yu song mo yi min shuo 現代粵語源於宋末移民說 [Modern Yue comes from the migration in the late Song dynasty – a hypothesis]. In Lau, C., *Xianggang ke yue fang yan bi jiao yan jiu 香港客粵方言比較研究 [A Comparative Study of Hakka and Yue in Hong Kong]*, 188-200. Guangzhou: Jinan University Press.

Luo, K. (2004). Yue yu xing cheng yu gu guangxin – jian tan yue yu de wen hua jia zhi he bao hu wen ti 粵語形成于古廣信——兼談粵語的文化價值和保護問題 [The Formation of Yue in Ancient Guangxin – also on the Cultural Value of Yue and its Protection]. In *Lingnan wen shi 嶺南文史 [Lingnan Culture and History]*, 2004(03): 20-22.

Luo, Y. (2013). *Ao men yue yu yin xi de li shi bian qian ji qi cheng yin 澳門粵語音系的歷史及其變遷 [Historical changes in Macao Cantonese Phonology and their Causes]*. Beijing: Beijing University. [PhD dissertation, Beijing University. Available at https://tra-oversea-cnki-net.ezproxy.is.ed.ac.uk/KCMS/detail/detail.aspx?dbcode=CDFD&dbname=CDFD1214&filename=1013208914.nh&uid=WEEvREcwSlJHSldRa1FhdkJtNEY_wQTQvUDM4U3BkK0w3SjZkUmVTMFVoWT0=$9A4hF_YAuvQ5obgVAqNKPCYcEjKensW4IQMovwHtwkF4VYPoHbKxJw!!&v=MjIwNzZWcnJLVkYyNkhiRzRGdGpOcTVFYlBJUjhlWDFMdXhZUzdEaDFUM3FUcldNMUZyQ1VSN3FmWU9kdUZDbm0=]

Mai, Y. (1993). Yue yu shih an yu de yi zhi fang yan 粵語是一支漢語的方言 [Yue is a dialect of Hanyu/ Chinese]. In *Yu wen jian she tong xun 語文建設通訊 [Language construction newsletter], Issue 41*. Cited in Mai, Y. (2009). Cong yue yu de chan sheng he fa zhan kan han yu fang yan xing cheng de mo shi 從粵語的產生和發展看漢語方言形成的模式 [Looking at the Formation Modes of Chinese Dialects from the Origin and Development of Yue].

Mai, Y. (1997). Zhong gu jing zu zi zai yue yu zhu zi fang yan de bu tong du yin ji qi li shih an yi 中古精組字在粵語諸方言的不同讀音及其歷史含義 [The pronunciation differences of Middle Chinese Jing series words in Yue dialects and its historical implications]. In *Journal of Chinese Linguistics*, Vol 25 (2). Cited in Mai, Y. (2009). Cong yue yu de chan sheng he fa zhan kan han yu fang yan xing cheng de mo shi 從粵語的產生和發展看漢語方言形成的模式 [Looking at the Formation Modes of Chinese Dialects from the Origin and Development of Yue].

Mai, Y. (2009). Cong yue yu de chan sheng he fa zhan kan han yu fang yan xing cheng de mo shi 從粵語的產生和發展看漢語方言形成的模式 [Looking at the Formation Modes of Chinese Dialects from the Origin and Development of Yue]. In *Fangyan 方言 [Dialect]*, 31(03): 219-232.

Nerbonne, J. & Heeringa, W. (2010). Measuring Dialect Differences. In: Schmidt, E. J. & Auer, P. (eds.) *Language and Space: Theories and Methods in series Handbooks of Linguistics and Communication Science*, vol. 1, 550-567. Berlin: Mouton De Gruyter.

Nerbonne, J. (2010a). Mapping Aggregate Variation. In Lameli, A., Kehrein, R. & Rabanus, S. (eds) *Language and Space. An International Handbook of Linguistic Variation*, vol. 2, 476-495. Berlin: Mouton De Gruyter.

Nerbonne, J., & Wieling, M. (2018). Statistics for Aggregate Variationist Analyses. In Boberg, C., Nerbonne, J. & Watt, D. (eds) *The Handbook of Dialectology*, 400-414. Hoboken, NJ: John Wiley & Sons.

Nerbonne, J., Colen, R., Gooskens, C., Kleiweg, P. & Leinonen, T. (2011). Gabmap — A Web Application for Dialectology. In *Dialectologia*, Special Issue II, 65-89. [Retrieved in Jan, 2020, from https://pdfs.semanticscholar.org/b4f8/2c72975bd2a79fc2d6cec2671fb9f86c67a6.pdf]

Orton, H. (1962). *Survey of English Dialects: Introduction*. Leeds: E. J. Arnold.

QGIS Development Team (2020). *QGIS Geographic Information System. Open Source Geospatial Foundation Project*. http://qgis.osgeo.org

Shao, H. (2016). *Yue xi zhan mao di qu yue yu yu yin yan jiu 粵西湛茂地區粵語語音研究 [The Phonological Study of the Yue Dialects spoken in the Zhan-Mao area in Western Guangdong]*. Guangzhou: Sun Yat-Sen University Press.

Sun, C. (2018). *Gu han yu yu yin 古漢語語音 [Phonology of Ancient Chinese]*. Shanghai: Shanghai Guji Press.

Tang, C., & Van Heuven, V. J. (2015). Predicting mutual intelligibility of Chinese dialects from multiple objective linguistic distance measures. *Linguistics*, 53(2): 285-312.

The Chinese University of Hong Kong. The Research Centre for Humanities Computing. (CUHK Humanum) (2020). *Multi-function Chinese Character Database*. [Retrieved in May, 2020, from https://humanum.arts.cuhk.edu.hk/Lexis/lexi-mf/search.php?word=%E6%A0%B8]

Therese Leinonen, Çagri Çöltekin, and John Nerbonne (2016). Using Gabmap. In *Lingua*, 178, 71-83.

Trudgill, P. (1983). *On dialect: Social and geographical perspectives*. New York: New York University Press.

Tung, T. (1968). *Han yu yin yun xue 漢語音韻學 [Chinese Phonology]*. Taipei: Wenshijie Publishing. Cited in Zhu, J. (2016). *Sheng yun xue 聲韻學 [Phonology (2nd Edition)]*. Taipei: Wunan Publishing Company Ltd.

Wang, L. (1936). *Zhongguo yin yun xue 中國音韻學 [Chinese Phonology]*. Commercial Press. Cited in Yue-Hashimoto, A. (1991). Yue yu fang yan fen qu wen ti chu tan 粵語方言分區問題初探 [On the Problem of Yue Dialect Classification]. In *Fangyan 方言 [Dialect]*, 3(03): 164-181.

Wang, L. (1985). *Han yu yu yin shi 漢語語音史 [A Phonological History of Chinese]*. Beijing: Chinese Academy of Social Science Press. Cited in Zhu, J. (2016). *Sheng yun xue 聲韻學 [Phonology (2nd Edition)]*. Taipei: Wunan Publishing Company Ltd.

Wang, L., & Qian, S. (1949). Dongguan fang yin 東莞方音 [Dongguan Phonology]. *Lingnan xue bao 嶺南學報 [Lingnan Journal]*, 10(1): 119-150.

Wang, L., & Qian, S. (1950b). Taishan fang yin 台山方音 [Taishan Phonology]. *Lingnan xue bao 嶺南學報 [Lingnan Journal]*, 10(2): 67-104.

Wiegand, P. (2007). *Oxford Student's Atlas*. Oxford University Press.

Wu, S. (1997). Zhongguo yi min shi, di 4 juan, liao song jin yuan shi qi 中國移民史，第 4 卷，遼宋金元時期 [The Migration History of China: Part 4 The Liao, Song, Jin, Yuan Periods]. Fujian People's Press. Cited in Lau, C. (2001). Xian dai yue yu yuan yu song mo yi min shuo 現代粵語源於宋末移民說 [Modern Yue comes from the migration in the late Song dynasty – a hypothesis]. In Lau, C.'s *Xianggang ke yue fang yan bi jiao yan jiu 香港客粵方言比較研究 [A Comparative Study of Hakka and Yue in Hong Kong]*, 188-200. Guangzhou: Jinan University Press.

Wu, W. & Zhan, B. (2008). Guangdong sheng de han yu fang yan 廣東省的漢語方言 [Chinese Dialects in the Guangdong Province]. In *Fangyan 方言 [Dialect]*, 30(2): 109-116.

Wu, W. (2012). Yue yu 粵語 [Yue]. In Zhongguo she hui ke xue yuan. Yu yan yan jiu suo 中國社會科學院語言研究所 [Chinese Academy of Social Science. Research Institute of Language (CASS)] & City University of Hong Kong. Yu yan zi xun ke xue yan jiu zhong xin (CityU LISRC) 語言資訊科學研究中心 [Language Information Sciences Research Centre], editorial body. (2012). *Zhongguo yu yan di tu ji : Di 2 ban. Han yu fang yan juan 中國語言地圖集 (第二版): 漢語方言卷. [Language atlas of China (2nd Edition): Chinese Dialects],* 125-133. Beijing: Shang wu yin shu guan.

Wurm, S. A., T'sou, B., Bradley, D., Li, R., Xiong, Z. & Zhang, Z. (1989). *Language Atlas of China* (1st Edition). Canberra: Australian Academy of the Humanities and the Chinese Academy of Social Sciences & Department of Linguistics, Research School of Pacific Studies, ANU. Hong Kong: Longman Group (Far East).

Xiong, Z. (1987). Guangdong fang yan de fen qu 廣東方言的分區 [Dialect classification in Guangdong]. In *Fangyan 方言 [Dialect]*, 3(03): 161-165.

Yan, M. M. (2006). *Introduction to Chinese Dialectology*. Munich: Lincom.

Ye, G. & Luo, K. (1995). Yue yu yuan liu kao 粵語源流考 [On the Origin of Yue]. In *Yu yan yan jiu 語言研究 [Linguistics study]*, 1995(01): 156-160.

You, R. (2016). *Han yu fang yan xue jiao qing 漢語方言學教程 [Courses in Chinese Dialectology]*. Shanghai: Shanghai Education Press.

You, R. (2018). *Han yu fang yan xue jiao qing 漢語方言學教程 [A Course in Chinese Dialectology]*. Shanghai: Shanghai Education Press.

Yuan, J. (1960). *Han yu fang yan gai yao 漢語方言概要 [An Outline of the Chinese Dialects]*. Beijing: Wen zi gai ge Press.

Yuan, J. (2001). *Han yu fang yan gai yao (di 2 ban) 漢語方言概要 (第二版) [An Outline of the Chinese Dialects (2nd Edition)]*. Beijing: Yuwen Publishing.

Yue-Hashimoto, A. (1991). Yue yu fang yan fen qu wen ti chu tan 粵語方言分區問題初探 [On the Problem of Yue Dialect Classification]. In *Fangyan 方言 [Dialect]*, 3(03): 164-181.

Zhan, B. & Cheung, Y. (1987). *A Survey of Dialects in the Pearl River Delta, Vol. 1, Comparative Morpheme-Syllabary*. People's Publishing House of Guangdong.

Zhan, B. & Cheung, Y. (1988). *A Survey of Dialects in the Pearl River Delta, Vol. 2, Comparative Lexicon*. People's Publishing House of Guangdong.

Zhan, B. & Cheung, Y. (1990). *A Survey of Dialects in the Pearl River Delta, Vol. 3, A Synthetic View*. People's Publishing House of Guangdong.

Zhan, B. & Cheung, Y. (1994). *A Survey of Yue Dialects in North Guangdong*. Jinan University Press.

Zhan, B. & Cheung, Y. (1998). *A Survey of Yue Dialects in West Guangdong*. Jinan University Press.

Zhan, B. (1981). Xian dai han yu fang yan 現代漢語方言 [Modern Chinese Dialects]. Hubei People's Press.

Zhan, B. (ed.) (2002). *Guangdong yue fang yan gai yao 廣東粵方言概要 [An Introduction to the Yue dialects in Guangdong]*. Guangzhou: Jinan University Press.

Zhongguo she hui ke xue yuan. Yu yan yan jiu suo. 中國社會科學院語言研究所 [Chinese Academy of Social Science. Research Institute of Language.] (1983). *Fang yan diao cha zi biao 方言調查字表 [The Dialect Survey Wordlist]*. Beijing: Shang wu yin shu guan : Xin hua shu dian Beijing fa xing suo fa xing.

Zhongguo she hui ke xue yuan. Yu yan yan jiu suo 中國社會科學院語言研究所 [Chinese Academy of Social Science. Research Institute of Language (CASS)] & City University of Hong Kong. Yu yan zi xun ke xue yan jiu zhong xin (CityU LISRC) 語言資訊科學研究中心 [Language Information Sciences Research Centre], editorial body. (2012). *Zhongguo yu yan di tu ji : Di 2 ban. Han yu fang yan juan 中國語言地圖集 (第二版): 漢語方言卷. [Language atlas of China (2nd Edition): Chinese Dialects]*. Beijing: Shang wu yin shu guan.

Zhou, Z. & You, R. (2019). *Fang yan yu Zhongguo wen hua 方言與中國文化 [Dialects and the Chinese Culture]*. Shanghai: Shanghai People's Press.

Zhou, Z. (1966). Qieyun de xing zhi he ta de yin xi ji zhu 切韻的性質和它的音系基礎 [The Nature of Qieyun and its phonological basis]. In Zhou, Z., *Wen xue ji 問學集*, 434-473. Beijing: Zhonghua Book Company.

Zhu, J. (2016). *Sheng yun xue (di 2 ban) 聲韻學 (第二版) [Phonology (2nd Edition)]*. Taipei: Wunan Publishing Company Ltd.