Common Yue:
A Comparative Study of
Yue Dialect Historical Phonology

John Carlyle

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Arts

University of Washington

2020

Committee:

Zev Handel

William Boltz

Program Authorized to Offer Degree:
Asian Languages and Literature

University of Washington

Abstract

Common Yue:
A Comparative Study of
Yue Dialect Historical Phonology

John Carlyle

Zev Handel
Asian Languages and Literature

Common Yue is a comparative study of the phonology of the Yuè dialects of Guǎngdōng and Guǎngxī. It draws upon the lexical studies and *zìyīn* 字音 character reading data of 21 Yuè dialects as recorded in the fieldwork of Zhān, Cheung, et al. (1987a,b; 1994; 1998), Xiè (2007), and Lǐ (2014) and Xiǎn (2016).

Common Yue differs from previous studies in that: 1) Labiodentals are separate from velar fricatives. Labiodentals were historically a secondary class of phonemes used to approximate northern *koiné* loans of late origin. 2) Sonorant initials occur in both the *yīn* 陰 and *yáng* 陽 registers of tone, as is common in Mín, Hakka, and Shē dialects. 3) Labiovelars are phonemic units rather than sequences of initial and medial segments. Remaining "medial" [i] is treated as a fused part of vocalic clusters rather than an independent segment due to its limited distribution, i.e. it only occurs before the phonetic segments [a] and [ɔ]. 4) The split of the historical *yīnrù* 陰入 tone into high and low reg-

isters was conditioned by the feature [open] in vocalic nuclei. Vowel chain shifts obscure this original conditioning environment, but it is still partially observable in conservative dialects. 5) Common Yue has two reflexes of tone 4, tone 4a and tone 4b. Tone 4a is reconstructed where Sìyì dialects have tone 1 and other dialects have tone 4, while tone 4b is reconstructed where all dialects have tone 4. The latter tone correlates with *koiné* readings of sonorant onset words and seems to be at least partially secondary in origin like the labiodentals.

These revisions allow for a more complete description of diachronic Yuè phonology and open the door to exploration of the genetic relationship of Yuè to the daughter dialects of Early Southern Highlands Chinese (Coblin, 2018) and other neighboring languages.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

Yuè 粵 dialects are made up of several lexical layers due to waves of Chinese immigration to the Lǐngnán 嶺南 region. While previous scholars' diachronic studies of Yuè dialect phonology stand as important contributions to the reconstructive enterprise, they predate a modern understanding of the methodological challenges lexical layering presents to the comparative study of Chinese dialects (McCoy, 1966; Tsuji, 1980; Yue-Hashimoto, 2000; 2006). Common Yue is an attempt to remedy previous deficiencies by accounting for lexical layering during the application of the comparative method. In Common Yue, we apply the comparative method in its strict sense to actual lexical cognates to establish correspondences, not character readings divorced from context. We believe that we may use these correspondences to seriate layering present in character readings. The result is the reconstructed phonology of 241 basic lexical items.

Common Yue draws upon the lexical studies and *zìyīn* 字音 character reading data of Zhān, Cheung, et al. (1987a,b; 1994; 1998), Zhān et al. (2000), Yue-Hashimoto (2005), Xiè (2007), and Lǐ (2014) and Xiǎn (2016). Of these sources, Zhān, Cheung, et al. (1987a,b) and Xiè (2007) also include non-Yuè data in the Lǐngnán area. Before applying

the comparative method and reconstructing Common Yue, we examine all data and propose phonological and lexical criteria for the classification of Yuè as a dialect family. Of the dialects recorded in these studies, we classify 70 as Yuè dialects. We then propose an internal grouping by measuring the lexical distance between the 70 sites. Finally, we select 21 of these sites on the basis of lexical of diversity and apply the comparative method to reconstruct a phonological common system for the dialect family. The common system is intended to represent the theoretical phonological ancestor of all 70 sites, but we only present the correspondences of 21 as a matter of practicality. Shorter correspondences are easier to understand. Comparing closely related dialects yields correspondences that are largely uniform and obscures important differences. We occasionally refer to dialects outside the selected 21 where they help illustrate our proposed developments.

Common Yue differs from previous reconstructions of Yuè phonology in that: 1) Labiodentals are separate from velar fricatives. Labiodentals were historically a secondary class of phonemes used to approximate northern *koiné* loans of late origin. 2) Sonorant initials occur in both the *yīn* 陰 and *yáng* 陽 registers of tone, as is common in Mín 閩[1], Hakka, and Shē 畲 dialects. 3) Labiovelars are phonemic units rather than sequences of initial and medial segments. The remaining "medial" [i] is treated as a fused part of vocalic clusters rather than an independent segment due to its limited distribution, i.e., it only occurs before the phonetic segments [a] and [ɔ]. 4) The split of the historical *yīnrù* 陰入 tone into high and low registers was conditioned by the feature [open] in vocalic nuclei.

[1]The citation tone for this word is usually tone 3, Mǐn, but Norman noted that the native Mín pronunciation takes the tone that corresponds historically to Mandarin tone 2.

Vowel chain shifts obscure this original conditioning environment, but it is still partially observable in conservative dialects.

These revisions allow for a more complete description of diachronic Yuè phonology and open the door to exploration of the genetic relationship of Yuè to the daughter dialects of Early Southern Highlands Chinese and other neighboring languages (Coblin, 2018).

1.1 Overview

The Yuè dialects are spoken throughout the Pearl River Basin in Guǎngdōng 廣東 and Guǎngxī 廣西 provinces, extending to the extremes of the basin's tributaries on the East River (*Dōngjiāng* 東江) and the North River (*Běijiāng* 北江). The extent to which the Yuè dialects are distributed beyond the West River (*Xījiāng* 西江) is controversial and largely depends upon the classification of Yuè and Pínghuà (*Pínghuà* 平話) dialects. Although the western dialects have received better documentation in recent years, thanks to the efforts of scholars like Lǐ (2000) and Xiè (2007), their boundaries and exact relationship to the Yuè dialects remain unsettled. We take the view that Southern Pínghuà dialects are actually Yuè dialects and that Pínghuà does not constitute a valid classification for all the dialects it is currently proposed to include. We justify this position with new evidence presented in Chapter 2.

1.1.1 External Classification

Traditionally, Chinese dialect families have been classified primarily according to the development of historical voiced obstruent initials in various tone classes, the development of historical velar initials before high front vowels, and the retention of stop consonant endings (Li, 1973; Ting, 1982). These sorts of phonological criteria present problems for classifying Yuè dialects. There is no unified development of voiced obstruent initials across tone classes in Yuè dialects. The dialects around the Pearl River Delta demonstrate a trend in which historically voiced stops became aspirates in the Píng (平)and Shàng (上) tone classes, but plain stops in the Qù (去) and Rù (入) tone classes. In the western region, they tended to evolve entirely into plain stops. In certain parts of the the southwest region, they tended to involve entirely into aspirated stops. Most Yuè dialects do not palatalize velar initials before high front vowels nor lose historical stop endings, but this is insufficient to distinguish them from neighboring southern dialect families like Hakka that demonstrate the same trends.

Norman (1988) suggested that the Yuè dialects can be distinguished on the basis of a split of the historically voiceless *yīn* (陰) register of the Rù tone class and unique lexical items. This distinguishes them neatly from all other large Chinese dialect families except the controversial Pínghuà family. A consensus among Yuè scholars is that the southern group of the proposed Pínghuà family should be grouped with other Yuè dialects. The northern group eludes definite classification but exhibits curious phonological tendencies shared with Southern Xiāng 湘 and Northern Patois (*tǔhuà* 土話). Before we begin our

comparison, we will pursue this issue further. It is critical that the Yuè family be defined explicitly by some criteria so we have a basis upon which we can decide which dialects to exclude from our study.

1.1.2 Internal Classification

Early attempts to internally classify Yuè dialects were informed by geography and simple listings of linguistics features without further discussion of why these features guarantee or disqualify membership to certain groups.

Yuán et al. (1983) attempted the first internal classification of Yuè dialects based entirely on geographic distribution. They establish five groups, 1. Yuèhǎi 粵海, including the majority of the region around the Pearl River Delta and the West River 2. Qīnlián 欽廉, including the regions around Qīnzhōu 欽州 and Liánzhōu 廉州 3. Gāoléi 高雷, including the region around Gāozhōu 高州 and Léizhōu 雷州 4. Sìyì 四邑, including the area around Xīnhuì 新會, Táishān 台山, Ēnpíng 恩平, and Kāipíng 開平 5. Guìnán 桂南, including the southern part of Guǎngxī around Wúzhōu 梧州, Róngxiàn 榮縣, Yùlín 鬱林, and Bóbái 博白 . Zhān (1981) listed linguistic features of each of these regions. He did not attempt to justify the groups in terms of the features, but rather listed the features post-hoc.

Xióng (1987) separated Yuè dialects into five groups based on three phonological criteria. The first criterion is the development of historical voiced stops and affricate initials. Dialects in which these initials unconditionally evolved into plain stops and affricates, he groups in to the Gōulòu 勾漏 group, a name he adopted from Yáng et al. (1985)'s con-

ventions. Dialects in which these initials unconditionally evolved into aspirated stops, he grouped in to the Wúhuà 吳化 group. The remaining dialects demonstrate a trend of these initials evolving into aspirates in the Píng and Shàng tone classes and plain stops and affricates in the Qù and Rù tone classes. The second criterion is the development of QYS *th-* initial (*tòumǔ* 透母).[2] Dialects in which this initial evolved into /h-/, he put in to the Sìyì group. Finally, he separated the last two groups on the criterion of the development of the QYS *s-* (*xīnmǔ* 心母)initial. Dialects in which it evolved into /ɬ-/, he classified as members of the Gāoyáng 高陽 group. All remaining dialects he relegated to the Guǎngfǔ 廣府 group.

Yue-Hashimoto (1988; 1991) criticized Xióng's grouping on the grounds that it does not address the dialects of Guǎngxī and that it uses only a limited set of criteria with dubious explanatory power. She notes that Xióng's second criterion is met in the phonology of the Nánhǎi 南海 dialect, but he does not classify it as a Sìyì dialect, and his third criterion is met in the phonologies of the Táishān and Kāipíng dialects but he does not classify them as Gāoyáng dialects. She concludes that "the use of a single criterion not correlated with any others in classifying the Yuè dialects, even [just] in Guǎngdōng seems to be unable to reflect the actual linguistic situation" (Yue-Hashimoto, 1988, p.13).

Yue-Hashimoto (1988; 1991) opted for a more comprehensive approach grounded in the traditional methodology of linguistic geography. She selected 31 phonological

[2]The *Qièyùn* System (QYS) is the phonological system derived for the *f ǎnqiè* 反切 glosses spellings recorded in the rime dictionary *Qièyùn* by means of traditional philological methods complemented by basic phonological comparison. This system is often called "Middle Chinese" or "Ancient Chinese" by other scholars. All QYS transcriptions are given in italics using Baxter (1992)'s transcription.

features and 3 lexical features as isoglosses and studied their geographical distribution across 47 Yuè dialects. She observed that the greatest concentration of isoglosses falls on the boundary between the region encompassing the Sìyì, Yángchūn 陽春 and Yángjiāng 陽江 districts and the rest of the Pearl River Basin. Therefore, she divided the Yuè dialects into two macro-groups, the Sìyì-Liǎngyáng 四邑兩陽 macro-group and the Pearl River Delta macro-group (*Sānjiǎozhōu qū* 三角洲區).[3]

She divided the former macro-group into the Sìyì and Liǎngyáng groups. The Sìyì group is characterized by the merger of the *yīn* register, Píng and Qù classes (Isogloss T2a), the development of QYS *th-* into the initial /h-/ (Isogloss I6b), and the nondistinction of the vowel nuclei /a/ and /ɐ/ (Isogloss F1b). The Liǎngyáng 兩陽 group is characterized by a higher pitch value for *yáng* register tones than *yīn* register tones (Isogloss T1) and contour tone values for modern Rù class tones (Isogloss T9).

She divided the Pearl River Delta macrogroup into three groups, the Northern Delta (ND) group, the Southern Delta (SD) group, and the Guǎngfǔ (GF) group.

She characterized the ND group by its development of historical voiced stops and affricates into plain stops and affricates (Isogloss I1), although exceptions apply to the Nánhǎi 南海, Huàxiàn 化縣, and Wúchuān 吳川 dialects. She observed that this development only obtains in the colloquial stratum of the Shùndé 順德, Gāoyáng 高陽, and Gāomíng 高明 dialects.

---

[3]In the original text, Yue-Hashimoto used the terms “group” (qū 區), “subgroup” (piān 片), and “subdialect group” (xiǎopiān 小片). To reduce the confusion, ambiguous hierarchical terms such as “(dialect) subgroup” and “subdialect group” might cause, I have relabelled her hierarchy in English as “macro-group,” “group,” and “subgroup” respectively.

She divided the ND group into the Sānyì-Zhàoqìng subgroup (SZ) and the Inland subgroup (Inld). The SZ subgroup is characterized by a tendency of the rime table rime groups *Xiāo* 蕭, *Xián* 咸, and *Shān* 山 of divisions II and IV to merge in colloquial lexicon (Isogloss F12)[4], palatalization of velars before inner rime table rime groups including division III of *Líu* 流, *Shēn* 深, and *Zhēn* 真 (Isogloss I3a)[5], the development of the QYS initials *j-* and *hj-* into /h-/ (Isogloss I8), and the use of /lei/ as a proximate deictic (Isogloss VIc). The Inld group she characterized by occasional development of plain bilabial and dental stop initials into voiced or imploded initials (Isogloss I7) and the development of alveolar affricates into dental stops (Isogloss I5a).

She divided the SD dialect group into the Qīnlián and Zhōngshān 中山 (ZS) subgroups. The ZS subgroup, she argues, is characterized by having just six tones due to a merger of the *yīn* and *yáng* registers of both the Shàng and Qù tone classes. She lists many isoglosses to characterize the Qīnlián group, but perplexingly none of them apply consistently to modern descriptions of the dialects in this area. Merger of finals /-u/ and /-y/ (Isogloss F6) does not obtain in Xiè (2007)'s description of the Qīnzhōu and Běihǎi dialects, nor does the merger of the *yīn* and *yáng* registers of the Shàng and Qù tone classes (Isogloss T3).

The remaining dialects she grouped together as the GF group and characterized them

---

[4]This essentially equates to instances of the finals /ɛw, ɛm, ɛn/ occurring in the II and IV divisions of the rime table finals given. Instances of this are marginal in practice and it is misleading to qualify this as a "merger."

[5]This coincides with syllables in which we reconstruct the vowel nucleus /i/ with codas that involve closure at the front of the oral cavity in Common Yue. The actual distribution of this feature is limited and extends past the boundaries of the SZ group as defined by her in this article.

by their typical nine tone system with rising Shàng, level Qù and Rù tone classes, and all *yīn* register tones being higher if fundamental frequency than *yáng* register tones.

In her reconstruction of proto-Yuè initials, Yue-Hashimoto (2006) presents some minor changes to her earlier system of classification, but does not explain the criteria these revisions are based upon. We shall attempt to infer these criteria based on the isoglosses of her previous studies. First, she renames Sìyì group to the Wǔyì 五邑 (WY) group, and by extension, the Sìyì-Liǎngyáng macro-group to the Wǔyì-Liǎngyáng 五邑兩陽 (WL) macro-group, to conform with the modern administrative term for the region. She further subdivides the WY group into the Xīn'ēn 新恩 (XE) and Kāihè 開鶴 (KH) subgroups, presumably on the bases of /p-/ evolving to /v-/ and /p^h-/ evolving to /h-/ (Isogloss I2) in the Kāipíng and Hèshān 鶴山 dialects. She rebrands the previous Qīnzhōu subgroup as the Guǎnlián 莞廉 (GL) subgroup and creates a new subgroup in the SD group she dubs the Guǎnlián Inland (GL Inld) group. The latter includes many dialects she previously classified as Inld dialects. Presumably, the basis for this new grouping is the evolution of the initial /s-/ into /ɬ-/ (Isogloss I10), with the caveat that historical /s-/ and /ʃ-/ remain unmerged, and the occasional development of implosive initials (Isogloss I5a). Of her new list of ND Inld and GL Inld dialects, only the Téngxiàn 藤縣 dialect bears both of the above traits yet is still placed in the ND Inld group. This might be a typographical error. Finally, she divides the GF group into Central and Inland subgroups. This seems to be informed by the preservation of the distinction between the modern vowel nuclei /ɐ/ and /ø ~ɵ ~œ/ (Isogloss F2c).

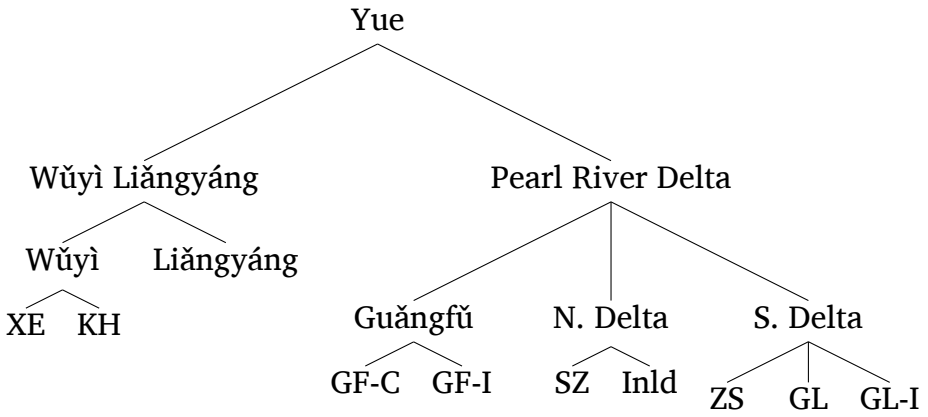


Figure 1.1: Yue-Hashimoto (2006)'s Classification of Yuè Dialects

In the same year, Kwok (2006) criticized Yue's previous attempts at classification for lacking diachronic bases. He raises some diachronic concerns about classification involving tonal development, but does not produce a complete new system of classification. While diachronic bases for classification are certainly valid, they do not invalidate classifications derived from the traditional geographical method. Dialects rarely develop in isolation. Changes are often better understood in terms of a wave model rather than a tree model. The geographical method accounts for the geographical distribution of linguistic features that impact speaker perceptions of language distance and mutual intelligibility. Yue-Hashimoto carefully selected her isoglosses based on years of first-hand fieldwork. No doubt the perceptual differences she uncovered with these isoglosses are valid even if they do not all owe their origin to a clear branching diachronic origin.

At the time of writing, Yue's study is the only complete, rigorous attempt to classify all Yue dialects according to accepted linguistic methodology. The fact that it does not

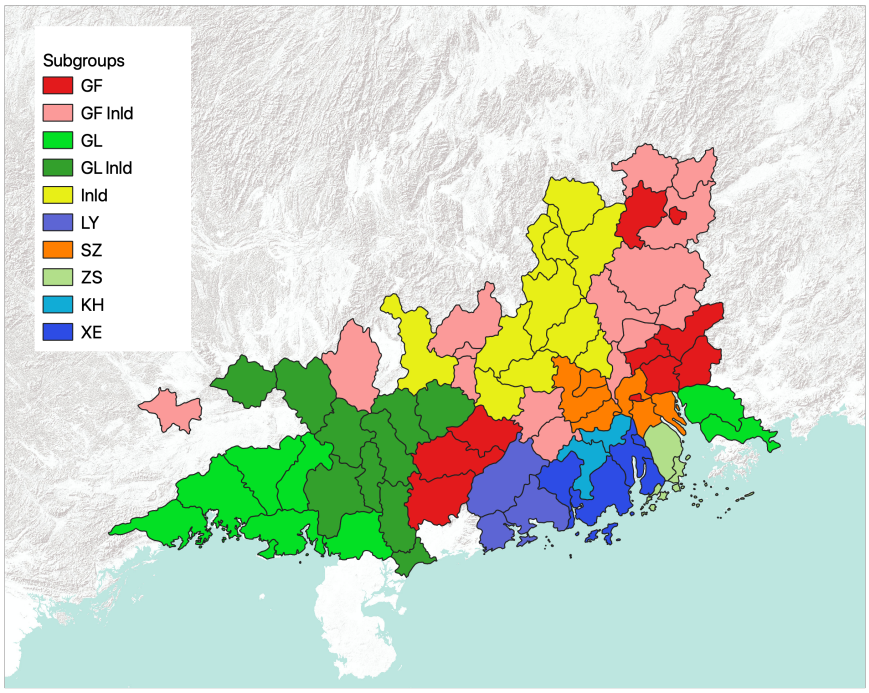


Figure 1.2: The Distribution of Yue-Hashimoto (2006)'s Dialect Groups

rely on diachronic analysis does not invalidate it. However, the geographic method has its limitations. Drawing isoglosses by hand is a laborious process prone to error. Investigators have to restrict the number of features they choose as criteria for isoglosses as a matter of practicality. In Yue's particular case, she also chose to characterize many of her phonological isoglosses in terms of QYS distinctions rather than discrete observable differences in phonology. It biases her investigation towards older changes and brings up unnecessary diachronic connotations. In the following chapter we will undertake a new

internal classification using techniques from the subfield of dialectometry. Dialectometry seeks to quantify differences across dialects in an objective way. It largely relieves investigators of the burden of arbitrarily choosing features to compare and, coupled with the power of modern computing, offers new more precise ways to represent dialect diversity visually on maps.

1.1.3 The Pínghuà Controversy

Zhāng (1982) first used the term Pínghuà as a unique classification in his study of Chinese loan words in Zhuang languages. He adopted the term from the Guǎngxī indigenous word for Chinese varieties that are not Mandarin, Guǎngfǔ Yuè or Hakka. The term Pínghuà contrasts with the indigenous word "báihuà 白話" which he explicitly states refers to Guǎngfǔ Yuè, *not necessarily* other varieties of Yuè. He lists Pínghuà's defining phonological characteristics as the unconditional evolution of historical voiced obstruents into plain stops, the evolution of the historical initial /s-/ into /ɬ-/, and the split of the historically voiced *yáng* 陽 register Rù tone into two separate tone classes. The first two characteristics are identical to Xióng (1987)'s criteria for Gōulòu group membership.

Zhāng split the Pínghuà dialects into a southern group (*Guìnán Piān* 桂南片)and a northern group (*Guìběi Piān* 桂北片). He described the southern group as being located in Bīnyáng 賓陽 and Héng 橫 counties, the Yōngníng 邕寧 district of Nánníng 南寧 city, and the outlying rural regions around the Western river. The northern group, he claimed, is located in Língchuān 靈川, Yǒngfú 永福, Luóchéng 羅城 and Róngshǔi 融水 counties as

well as in the Línguì 臨桂 district of Guìlín 桂林 city. He noted that people of the southern group *do not use the word Pínghuà when referring to their own language*, but rather mostly use toponyms as linguistic endonyms.

At the time of publication, Zhāng specialized in the study of Zhuang language, not Yuè dialects. It is not clear if he was aware that many western Yuè dialects, belonging to Yue-Hashimoto's Inland and Guanlian Inland subgroups or the Gōulòu group share many of the same phonological features of his proposed Pínghuà group and that the Bīngyáng dialect had already been classified as Yuè (Tsuji, 1980). Even if Zhāng was aware, he did not adopt a stance on Pínghuà dialects' relationship with Yuè dialects as a whole but only on their relationship with the Guǎngfǔ group. He did not brand his phonological characteristics as necessary or sufficient conditions upon which one could classify a dialect as Pínghuà.

For better or worse, his new terminology stuck. Three years later, Yáng et al. (1985) listed Pínghuà as a group separate from the Yuè dialects, parroting Zhāng's phonological characteristics, but this time explicitly contrasting it with their newly coined Gōulòu group, despite listing near identical phonological characteristics. Curiously, Yáng et al. grouped the dialects of Héng County and Yōngníng district as their own subgroup belonging to Yuè. The only example of southern Pínghuà they give is Nánníng Tíngzǐ 亭子 dialect, leaving Bīnyáng's status ambiguous. In 1987, The Language Atlas of China used Zhāng's proposal to map Pínghuà as a group separate from Yuè on the same level as the seven other well-recognized Chinese dialect families (Wurm et al., 1987). This cast a

spotlight on the fledgling family and ignited a passionate debate among dialectologists. Ting (1996) argued that the scholars behind the elevation of Pínghuà to a major dialect family juxtaposed with Yuè and Xiāng in the atlas did not offer sufficient justification.

In the late 20th century and early 21st century, the classification of Pínghuà became a popular topic among Chinese scholars. The earlier literature tended to take an all or nothing approach. Some scholars defended Pínghuà as a separate group on the basis of a few unique lexical items and grammatical constructions, e.g., Wéi (1996). Others contended that all Pínghuà dialects are Yuè due to phonological and lexical similarity (Liáng, 1997; Zhān, 2001). Differences could be attributed to contact with Mandarin and other varieties of Chinese north of Guìlín. Finally, a consensus was formed that the the southern group of Pínghuà demonstrates enough commonality with Yuè dialects to warrant inclusion in the group, while the northern group of Pínghuà should be excluded (Qín, 2000; Wáng, 2001; Wǔ, 2001; Xiè, 2001; Zhān, Cuī, et al., 2003; Qín, 2007; Xiè, 2007). The northern group demonstrates more commonality with the unclassified Patois of the region both phonologically and lexically. Lǐ Liánjǐn stands out as a sole dissenter. In his survey of Pínghuà dialects, he included the Téngxiàn and Yùlín dialects, whose previous classification as Yuè was uncontroversial (Lǐ, 2000). Later, he went on to claim that all Gōulòu group Yuè dialects should be reclassified as Pínghuà (Lǐ, 2005).

The controversy never gained traction in English language scholarship, although it has been discussed (DeSousa, 2017). Tsuji (1980) had no qualms classifying the Guǎngxī dialects of his study, such as Bīnyáng dialect, as Yuè dialects a decade before the term

Pínghuà designation was coined. Yue-Hashimoto (1979) also classified the Téngxiàn dialect as a Yuè dialect without any hesitation nearly 30 years before Lǐ's survey. She never broached the issue of Pínghuà directly in her more recent studies of Yuè dialect historical phonology, but tellingly grouped all southern group Pínghuà as part of her Guǎnlián Inland dialect group in her reconstruction of proto-Yuè initials.

We follow the consensus and treat the southern group of Pínghuà as Yuè dialects, leaving the northern group for further study and comparison with Northern Guǎngdōng and Southern Húnán Patois.

1.2 Literature Review

The study of the comparative phonology of the Yuè dialects stands out as one of the earliest explorations of diachronic phonology in Chinese dialects. John McCoy undertook the first reconstruction of "Proto-Cantonese" in his 1966 dissertation using his own fieldwork on the Szeyap (Sìyì 四邑) dialects (McCoy, 1966). Tsuji Nobuhisa produced a similar dissertation in 1977 and reconstructed "Proto-Yue" using his fieldwork on the Yuè dialects of Guǎngxī 廣西 (Tsuji, 1980). Both scholars elicited their data and organized their correspondences sets using Qièyùn 切韻 System (QYS) phonological categories and character readings rather than actual words. They viewed the Yuè dialects as more or less direct descendants of "Ancient Chinese" (i.e., "Middle Chinese"), although McCoy did bring up issues of lexical layering in his discussion of elicitation. He claims to have recorded only colloquial forms by eliciting readings of characters in the *Fāngyán diàochá zìbiǎo* 《方言

調查字表》by means of colloquial words in which such “readings” occur. Unfortunately, this context is not supplied in his listing of the data. He rejected what he perceived to be “irregularities” in pronunciation by omitting elicitations that did not agree with the phonological categories of the rime tables (McCoy, 1966, p.21). This method is undesirable as it assumes a priori that all colloquial words in Yuè must have derived from a phonological system that agreed completely regularly with rime table categories. Close study of the lexicon of the Yuè dialects reveals this is not the case. Tsuji also uses the *Fāngyán diàochá zìbiǎo* as a guideline when presenting his data, but does not elaborate on any further details about how he elicited his forms (Tsuji, 1980, p.204).

Years later, Prof. Anne Yue-Hashimoto undertook a series of similar “Proto-Yue” studies using her own fieldwork and the fieldwork of Y.R. Chao. While she did not resort directly to QYS categories to establish correspondences, she still used character readings as the unit of comparison rather than words (Yue-Hashimoto, 2000; 2006). She believes that in addition to a “Pan-Han Stratum,” consisting of more literary northern *koiné* origin words, Yuè dialects also contain other colloquial substrates that predate these later introductions. Because of this, she argues that comparative Yuè phonology is best investigated on its own terms.

More recently, Karen Huang attempted to reconstruct the vowels of Proto-Yue, also nominally eschewing QYS categories, but explicitly states that she limited her data set to what she calls the “literary stratum” of Yue lexicon as it is reflected in “pronunciation of characters” (Huang, 2009). This qualification seems to be a disclaimer about the character

reading data she used in her reconstruction rather than an actual judgement of the forms she reconstructs. Naturally, she uses character readings as the unit of comparison. Lexical layering is not brought up again in her discussion, but rime table categories see reference.

Yue-Hashimoto does not reconstruct multilayered “readings” of characters to reflect lexical layering, even when such layering is clearly present in single dialect points. This omission ultimately hurts the explanatory power of many of her proposals. For example, she argues that the labiodental /f/ initial of modern Yue dialects must have had a secondary origin. She maintains that Proto-Yue bilabial stops developed into labiodental fricatives before high back rounded vowels (Yue-Hashimoto, 2006, p.99). However, many dialects reflect layered character readings containing both bilabial stops and labiodental fricatives, presumably deriving from separate words of differing historical origin. Such correspondences are erratic. The stop readings and fricative readings do not occur consistently in the same dialects across “cognate” readings. Many of the dialects have both readings, reflecting two different words written using a common character. It is not possible to predict which dialects retained the bilabial stop readings and which innovated the labiodental fricative readings consistently in her model. It ultimately is a matter of which word is being said that decides the initial of the reading of the character used to write that word and whether or not such words have been displaced among the dialects. If one uses words rather than character readings as the basis for establishing cognates and regular correspondence, it is impossible to suppose any sort of conditioned change of bilabial stops in Yuè dialects.

Yue-Hashimoto only reconstructs one character reading in these cases, favoring the presumably earlier bilabial stop reading. This choice is not compatible with the comparative method. Proto-language phonologies must be internally consistent because sound change is exceptionless. Exceptionless sound change guarantees consistent correspondences. The appropriate way to handle layered readings is to set up multiple forms in the common system that can account for different words, and by extension different "readings." In other words, multiple forms are set up to handle words that gave rise to modern "colloquial character readings" and "literary character readings."

Yue-Hashimoto also reconstructs complex consonant cluster onsets. Some of these are based on partially reduplicated or dimidiated forms, while others are based on mixed cognate sets. The former is adopted from a proposal by McCoy, and considerations of historical origin aside, are generally workable (McCoy, 1986). The latter is problematic. Just as with the bilabial plosives, these correspondences are not consistent. It is not possible to predict the development of consonant clusters in her proto-system on an individual dialect basis. We must reject such correspondence sets if we are to apply the comparative method *sensu stricto*. In these erratic cases, multiple common system forms must be set up to account for all the reflexes.

Recently, the issue of reconciling the comparative method with "stratigraphic" analysis of lexical layers in Chinese dialects has been a matter of on-going discussion among specialists of historical Chinese dialectology (Akitani and Handel, 2012; Chén, 2013; Handel, 2015). Generally, I adopt the stance of W. South Coblin, which itself is a synthesis

of the dialogue (Coblin, 2015, p.6):

1. The method of comparative reconstruction is essential for a clear understanding of the history of Chinese dialects. (Akitani & Handel)
2. Identification and seriation of lexical strata constitutes a fundamental part of that reconstructive enterprise. (Chén)
3. Analysis of lexical strata should be incorporated into the ongoing process of comparison itself, rather than separated from it and dealt with beforehand. (Coblin)

In more recent work, Coblin attempts to seriate layers as he establishes regular patterns of correspondence among character readings. While I do not think such a method is inherently invalid, it is nonetheless vulnerable to bias when particular correspondences are obscured by heavy influences of layering. For example, tone 4 in Common Yuè is difficult to reconstruct on the basis of character readings alone. "Pure" sets of correspondence that do not mix in any way with the tone 6 correspondence sets after obstruent initials in character reading data are exceptionally rare. This is because tone 4 and tone 6 had already merged in the northern *koiné* that influenced Yuè dialects in the 10th and 11th centuries after obstruent initials in the well-known *yángshàng biàn qù* 陽上變去 tone shift. As such, tone 4 is often replaced with newer tone 6 reflexes in character reading data. However, in the correspondence sets established for true lexical items, unambiguous sets suggesting tone 4 are much more readily represented. In extreme cases like this, it is desirable to find an objective way to establish "true" correspondence sets.

In cases where we are fortunate enough to have full lexical descriptions of dialects, we

should use them to establish correspondences. Incorporation of high-register vocabulary from a *koiné* and retention of lexicon from earlier substrates is much more apparent in the comparison of actual words. For any given cognate, usually dialects agree phonologically in regard to what layer it descends from. If they do not, usually the words in question consist of entirely different morphemes, making them easy to rule out as non-cognates. Character readings rob us of this context. In many cases, we cannot be certain how these readings were elicited. Therefore, we often cannot be sure if certain morphemes are cognate even if they superficially appear similar. The comparative method demands that we establish the matter we are comparing is cognate before setting up correspondences. Observing a shared lexicon with close meaning and regular phonological correspondence is the best way to establish cognacy. Comparing readings of Chinese characters alone is insufficient. Notes from field investigators about character readings, like traditional *wéndú* 文讀 "literary" and *báidú* 白讀 "colloquial" labels, can help diachronic investigators evaluate cognacy better. However, such notation alone is still too crude to guarantee cognacy. When Coblin compares character readings, he is aware that a common Chinese character does not guarantee cognacy, even if they occasionally share labels from field investigators. Only when such readings fall into regular phonological correspondences are they likely to be cognate. These observations are valuable for reconstructing the historical phonology of Chinese dialects, but we may reduce the error in these observations if we are able to use actual lexicon as the starting point for comparison.

1.3 Data

Aside from attempting to represent sufficiently diverse Yuè dialects, my primary consideration in selecting dialect sites is the status of their documentation. I searched for sites that had received comprehensive lexical documentation in addition to a comprehensive description of character readings. Yue-Hashimoto (2000; 2006) made use of a wider range of materials, often quoting multiple sources for single dialect points. However, many of these materials are either too limited to use in an exhaustive comparison or are inaccessible.

As a small part of her materials, Yue-Hashimoto included dialect descriptions of the late 19th century and early 20th century in her comparison. The documentation of Yuè dialects in this time is usually limited to short lists of character readings and remarks about how these dialects' phonologies compared to Standard Cantonese (Don, 1882; 1884; Ball, 1889; 1890; 1896; Saunders, 1896; Ball, 1900a,b). Morrison (1828) stands out as an exception. He provides an excellent documentation of the lexicon of the Guǎngzhōu 廣州 dialect of his time. However, he did not transcribe tones, thus limiting his description's utility. While these materials are interesting and valuable for understanding recent phonological developments in Yuè dialects, they are too limited to include in the current comparison. Yue-Hashimoto also made use of the fieldwork of McCoy (1966) and Tsuji (1980), but as noted earlier, neither of these investigators produced sufficient lexical documentation.

The first major part of Yue-Hashimoto's data is her own fieldwork and that of Y.R. Chao. Only her descriptions of the Téngxiàn 藤縣 dialect and the Táishān 臺山 dialect of Dàncūn 淡村 are published (Yue-Hashimoto, 1979; 2005). The former gives lexicon according to the traditional fashion. Characters and their readings are sorted into rime table categories. Yue-Hashimoto does go an extra step in supplying actual lexicon and phrases in which such readings occur, but there is no way to reference them efficiently in its published state. Without an index, not only would a reader have to be familiar with the rime table pronunciation of the words they are trying to reference, but also how such a word is said in the dialect in first place as well and the headword character under which it would occur. The latter gives a much more comprehensive lexical description and comes complete with a Standard Mandarin - English - Táishān glossary. We use this source in our analysis for the internal grouping of Yuè dialects, but do not include it in the Common Yue comparison since it is so similar to the Wǔyì dialects. The only published materials of Y.R. Chao are his brief descriptions of the Zhōngshān 中山 and Táishān dialects (Chao, 1948a,b). Like other short descriptions of the time, they are mainly concerned with how Yuè dialects realize rime table distinctions rather than the dialects' lexicons.

The next major part of her data is the fieldwork surveys of Zhān, Cheung, et al. (1987a,b; 1994; 1998). These surveys give 3,810 character readings and comparison of more than 1,200 lexical items with accompanying Standard Mandarin glosses for 45 Yuè dialects. They fit my criteria and make up the majority of the Common Yue data. The fieldwork was carried out by various teams of dialectologists with Zhān and Cheung

acting as editors and overseers of the investigation. Although some of the members of the teams changed between the three surveys, they for the most part stayed the same as did their convention of phonetic transcription. Sadly, the surveys do not make explicit which members of the team were responsible for the documentation of which sites. The studies include the names and demographic information of their informants. Usually, their is just one informant per site. Unfortunately, in the cases where there were two or more informants for a site, they do not make explicit in the listing of the data which forms were elicited from which informant nor how they reconciled any linguistic discrepancies between the informants in detail.

Yue-Hashimoto also references a description of the Nánníng Pínghuà 南寧平話 dialect of Tíngzǐ 亭子 found in Lǐ (2000). Lǐ's study covers equally long character reading descriptions as the Zhān and Cheung et al. studies for many so-called "Pínghuà" dialects, but does not include lexical description, so I have omitted it from comparison.

To these sources, I add Xiè (2007)'s large survey of dialects in Guǎngxī, Xiǎn (2016)'s master's thesis on the Yángjiāng 陽江 dialect, new descriptions of the Xìnyí 信宜 and Liánjiāng 廉江 dialect from Zhān et al. (2000), and Lǐ (2014)'s recent description of the Huàzhōu xiàjiāng 化州下江 and Wúchuān 吳川 dialects. These sources provide extensive character reading homophone lists and lexical documentation of their subjects. Conveniently, Xiè (2007)'s study even uses the same index of lexical forms as Zhān, Cheung, et al. (1987a), facilitating easy comparison. Xiè (2007) lists the names of his informants in a postscript thanking them, but does not provide more detailed information or even which

informant spoke which dialect. Zhān et al. (2000) is mainly a synthesis of his previous surveys he guided. The work, again, includes many investigators under his guidance, but does not specify their exact contributions. It also does not specify how the Xìnyí and Lián-jiāng descriptions were conducted. Xiǎn (2016) lists three informants and their relevant demographic information, but does not specify how she synthesized her description from these three sources nor which forms were elicited from which informants. Lǐ (2014) is a self-report. Lǐ is a native of the Wúhuà area and claims to be familiar with both the dialects he describes from his travels all over the area during the Cultural Revolution (Lǐ, 2014, p.1). Although this sort of report is not conventional fieldwork, we nonetheless take the study at face value for want of more rigorous descriptions.

Source	Citation
[1]	Zhān Bóhuì 詹伯慧, Cheung, Yat-shing 張日昇, et al., 1987b. *Zhūjiāng sānjiǎozhōu fāngyán zìyīn duìzhào* 珠江三角洲方言字音對照. Guǎngdōng Rénmīn Chūbǎnshè Zhān Bóhuì 詹伯慧, Cheung, Yat-shing 張日昇, et al., 1987a. *Zhūjiāng sānjiǎozhōu fāngyán cíhuì duìzhào* 珠江三角洲方言詞彙對照. Guǎngdōng Rénmīn Chūbǎnshè
[2]	Zhān Bóhuì 詹伯慧, Cheung, Yat-shing 張日昇, et al., 1994. *Yuèběi shí xiànshì yuè fāngyán diàochá bàogào* 粵北十縣市粵方言調查報告. Jìnán Dàxué Chūbǎnshè
[3]	Zhān Bóhuì 詹伯慧, Cheung, Yat-shing 張日昇, et al., 1998. *Yuèxī shí xiànshì yuè fāngyán diàochá bàogào* 粵西十縣市粵方言調查報告. Jìnán Dàxué Chūbǎnshè
[4]	Zhān Bóhuì 詹伯慧 et al., 2000. *Guǎngdōng yuè fāngyán gàiyào* 廣東粵方言概要. Jìnán dàxué chūbǎnshè
[5]	Yue-Hashimoto, Anne, 2005. *The dancun dialect of taishan*. Language Information Sciences Research Centre, City University of Hong Kong
[6]	Xiè Jiànyóu 謝建猷, 2007. *Guǎngxī hànyǔ fāngyán yánjiū* 廣西漢語方言研究. 2 vols. Guǎngxī rénmín chūbǎnshè
[7]	Lǐ Jiàn 李健, 2014. *Wú huà yuèyǔ yánjiū* 吳化粵語研究, Zhānjiāng shīfàndàxué zhōngguó yǔyánxué xuékē xīn shìyě xuéshù wéncóng. Zhōngguó shèhuìkēxué chūbǎnshè
[8]	Xiǎn Wéntíng 冼文婷, 2016. *Guǎngdōng yángjiānghuà yánjiū* 廣東陽江話研究. Master's thesis. Guǎngxī dàxué

Table 1.1: Dialect Studies Consulted in this Investigation

1.4 Methodology

I establish correspondences in Common Yue on the basis of regular sound correspondence found in comparanda made up of shared lexical items among the dialects. Some of these comparanda demonstrate phonological tendencies typical of the later northern *koiné*. In a few special circumstances, it is even possible to catch a glimpse of how Yuè dialects rendered northern phonemes they lacked in a more limited phonological system that predated the Common Yue system. For example, what are labiodental fricative initials in other Yuè dialects regularly correspond to aspirated bilabial stop initials in the NH dialect in words of suspected northern origin, such as the word for "steamed rice" /p^{h}an^{6}/ for /*fan^{6}/. Presumably, these stop initials were the native NH way of approximating northern /f-/, an independent phoneme they lacked, while other Yuè dialects adopted /f-/ directly. This correspondence is likely not the product of an actual sound change nor does it go back to an autochthonous phoneme in an earlier common phonological system. Despite this, we still "reconstruct" an individual phoneme in Common Yue to account for this regular correspondence.

This methodological consideration is necessary to prevent bias caused by seriation of lexical data before the comparative method is applied. In all of his diachronic studies of the Gàn, Hakka, and Shē dialects, Coblin asserts that he "hold[s] that the reconstruction exercise, as a methodology requires us to compare and reconstruct proto-forms for *all comparanda which show regular sound correspondences, regardless of their real or suspected*

age or origin. Where doublets or other such variants in cognate sets are present in individual dialects, and *comparable variational sets* are found from dialect to dialect, then we believe that such variations must be projected back to the common system. This will then result in a restored entity comprised of multiple lexical layers." Thus, we reconstruct phonemes in the common system for all regular correspondences across dialects even if these correspondences are not autochthonous to the phonological common system predating northern contact but are instead suspected to be products of Yuè dialects accommodating previously non-Yuè phonemes.

We adopt Coblin's methodology as defined above, but must add clarification to one point. Because we use words instead of character readings as our comparanda, we find cases of "doublets or other such variants in cognate sets" far less frequently. Although such variance exists, and it may even be possible to speculate about "variational sets" that derive from historical changes like the tone 4 and tone 6 *yángshàn biàn qù* merger, these sets are rarely "comparable." For example, in the etyma #046 MUD-EEL 鱔 /ʒian^{4a}/, #142 DRIED PERSIMMON 柿餅 /*ʒ[ʐ][^{4b}] piaŋ^{3}/, and #185 INSIPID 淡 /*dam^{4a} all dialects demonstrate regular tone correspondence for tone 4, except for the Sìyì dialects which generally have reflexes of tone 6, with a few exceptions. For etymon #142, the Hèshān 鶴山 dialect demonstrates cases of modern tone 4, which might either be a case of a *biànyīn* 變音 morphological changed tone or a case of GF lexical influence. For #185 INSIPID, the Kāipíng 開平 dialect has the regular reflex of tone 4a along with a regular aspirated initial. The Dàncūn 淡村 subdialect of Táishān also has this form for #185 IN-

SIPID, but not the Táishān dialect Zhān, Cheung, et al. (1987a) recorded (Yue-Hashimoto, 2005). These correspondences of tone 4 elsewhere to Sìyì tone 6 are rare. We handle them by splitting the etyma and reconstructing separate forms to account for the differing tone patterning, but in reality it is not possible to reconstruct these forms back to a unified ancestor *sensu stricto*. We do not handle these cases as "variational sets" separated by a slash as Coblin did because they are not sporadic variants one might find in character readings. Their geographical distribution suggests a complicated lexical history, so we, for the nonce, treat as separate etyma despite their apparent ultimate cognancy.

1.4.1 Methodological Considerations for Irregular Correspondence

Beyond the largely regular influence contact with northern *koiné* exerted on Yuè dialect lexicon, there are a large amount of irregular correspondence between Yue-Hashimoto's Northern Delta group and other Yuè dialects. The SZ subgroup demonstrates the greatest concentration of these irregular correspondences. It is not feasible to reconstruct separate proto-forms for all different combinations of the irregular correspondence without producing an improbably complicated common system, so they are not reflected in Common Yue.

Irregular sound change runs contrary to neogrammarian principles and remains a controversial topic in historical linguistics. Malkiel (1962; 1964) attempted to broach the issue of seemingly incomplete "weak" sound change in terms of dialect contact brought about by changing social circumstances. He argued that one dialect, if it gained social

prestige, could begin to influence the phonology of another dialect related if the speakers of the two speech communities had sufficient contact with each other. Sound changes that occurred in the prestige dialect could be replicated in the levelled dialect if the conditions for such change were still left intact.

Chen and Wang (1975) and Wang (1977) took a different tack to explain irregular sound change, arguing instead that sound change gradually permeates the lexicon on a word-by-word basis through a process called "lexical diffusion." Chen and Wang's earlier article only argues for the existence of lexical diffusion, presumably in tandem with neogrammarian regular sound change, while Wang's later article goes further to imply that lexical diffusion is the sole mechanism of sound change. He argues that the unit of sound change is the word rather than the phoneme and that what appears to be regular changes are actually cases of sound changes permeating all words in which the changed sound occurs.

Labov (1981) attempted to reframe the debate in response to Wang's provocative stance. He argued that the question should not be "regular sound change or lexical diffusion?" but rather "in what circumstances do regular sound change and lexical diffusion occur?" He set about trying to answer the latter question empirically over the last few decades (Labov, 1989; 1994; 2007; 2010; 2020). A simplified synthesis of his arguments is as follows:

1. Regular sound change is indeed regular as the neogrammarians first proposed rather the result of gradual fully permeated lexical diffusion. Once a regular sound change

begins, it obtains without exception in all applicable phonemic environments (Labov, 2020).

2. Transmission, characterized by structurally regular changes, seems to be correlated with child language acquisition, while diffusion, characterized by structurally irregular changes, seems to be correlated with adult language acquisition (Labov, 2007).

3. The scope of lexical diffusion is limited. It appears to be less common than regular sound change and is associated with complicated dialect contact scenarios (Labov, 1989; 2020).

A mature framework for handling irregular sound changes in diachronic comparison has eluded historical linguists for over a century and a half now. We will not attempt to propose one here as it is beyond the scope of the current study. We will simply point out where such diffusion happens without projecting it back to the common system since it is limited to only one specific region. However, we do wish to assert with certainty that the irregular changes in the Northern Delta group are probably instances of diffusion that was caused by GF speakers migrating west and inland in the last few centuries. We do not wish to take a stance on whether this diffusion was lexical or structural or draw any other binding conclusions on its precise motivations. We will attempt to accurately reflect all cases of irregularity so that other scholars can study the issue further.

1.4.2 Transcriptions Conventions

To facilitate easier comparison, I reanalyze the phonology of all dialects included in the comparison. The original descriptions of these dialects erred on the side of narrow transcription with a small amount of phonological reduction that was mainly based on traditional Chinese prescriptive categories of phonology. While these transcriptions mostly faithfully record observations made in the field and have value in understanding phonetic variation in Yuè dialects, they in many cases obscure regular sound correspondence. There are several cases of unresolved complementary distribution that only serve to make the correspondences seemingly more complicated. Comparing distinctive phonological categories makes it easier for both investigators and readers to understand the exact nature of the correspondence. Because of the large size of the data I am working with, I wrote a Python script to uniformly apply all changes in transcription to the data according to a set of rules I describe in Chapter 3.

Chapter 2

A NEW CLASSIFICATION

2.1 Classification as a Dialect Family

To begin our investigation of Yuè dialects, it is necessary to set down clear defining criteria for the family. Because of the sheer number of dialects in the Lǐngnán region, it is not practical to objectively derive these criteria using the fine-tuned methods we use for internal grouping. Instead, we propose nine diagnostic criteria that are generally thought of as unique characteristics of Yuè dialects. This rough method of classification is not meant to be the final say on the issue, but rather a convenient abstraction so we can narrow the scope of our investigation.

Two of our nine criteria are phonological, while the rest are lexical. They are as follows:

1. Presence of a phonemic lower *yīnrù* (陰入) tone, i.e. "Tone 9"

2. A distinction between open /a/ and more central /ɐ/

3. The word for "slaughter" takes the general form of /$t^{h}ɔŋ^{1}$/ (劏)

4. The word for "thing" takes the general form of /ɲɛ^{4}/ (嘢)

5. The word for “noon” involves the general form /an^{5}/ (晏)

6. Presence of a feminine suffix for animals that takes the general form of /na^{3}/ (嫲)

7. Presence of a 2nd person plural marker that takes the general form /ti^{6}/ (哋)

8. The word for child either takes the general form of /sɐj^{5} mɐn^{1} tsɐj^{3}/ (細民仔) or /sɐj^{5} lɔw^{3} kɔ^{1}/ (細佬哥)

9. The word for “(early) morning” involves some combination of the general forms /tʃiw^{1}/ (朝) and /tsɔw^{3}/ (早)

Table 2.1 demonstrates which criteria each dialect in our data meets. Dialects that are traditionally thought of as Pínghuà are marked with final _p, Hakka dialects are marked with _h, Mín dialects are marked with _m, Mandarin is marked with _g, Northern Guǎngdōng Patois are marked with _t, Bǎixìnghuà 百姓話 is marked with _b, Yǎhuà 伢話 is marked with _ng, and Xiāng 湘 dialects are marked with _x. We take any score 5/9 or higher to mean that a dialect site may be identified as a Yuè dialect for the sake of this study.

Table 2.1: Yue Dialect Family Classification

Abbr	Full Name	Src	1	2	3	4	5	6	7	8	9	Sum	Yue?
GZ	Guǎngzhōu (City) 廣州 (市區)	[1]	1	1	1	1	1	1	1	1	1	9	Yes
HK	Hongkong (City) 香港 (市區)	[1]	1	1	1	1	1	1	1	1	1	9	Yes

Table 2.1: Yue Dialect Family Classification

Abbr	Full Name	Src	1	2	3	4	5	6	7	8	9	Sum	Yue?
JT	Hongkong (Kamtin) 香港 (錦田)	[1]	1	0	1	1	1	1	1	1	0	7	Yes
MC	Macao (City) 澳門 (市區)	[1]	1	1	1	1	1	1	1	1	1	9	Yes
PY	Pānyù (Shìqiáo) 番禺 (市橋)	[1]	1	1	1	1	1	1	1	1	1	9	Yes
HXHS	Huāxiàn (Huāshān) 花縣 (花山)	[1]	1	1	1	1	1	1	1	1	1	9	Yes
CH	Cónghuà (City) 從化 (城內)	[1]	1	1	1	1	1	1	1	1	1	9	Yes
ZC	Zēngchéng (Cty. Seat) 增城 (縣城)	[1]	1	1	1	1	1	1	0	1	1	8	Yes
FS	Fóshān (City) 佛山 (市區)	[1]	1	1	1	1	1	1	1	1	1	9	Yes
NH	Nánshān (Shātóu) 南海 (沙頭)	[1]	1	1	1	1	1	1	1	1	1	9	Yes
SD	Shùndé (Dàliáng) 順德 (大良)	[1]	1	1	1	1	1	1	1	1	1	9	Yes
SS	Sānshuǐ (Southwest) 三水 (西南)	[1]	1	1	1	1	1	1	1	1	1	9	Yes
GM	Gāomíng (City) 高明 (明城)	[1]	1	1	1	1	1	1	1	1	0	8	Yes
ZS	Zhōngshān (Shíqí) 中山 (石岐)	[1]	0	1	1	1	1	1	1	1	0	7	Yes
ZH	Zhūhǎi (Qiánshān) 珠海 (前山)	[1]	0	1	1	1	1	1	1	1	1	8	Yes
DMSH	Dòumén (Shànghéng) 斗門 (上橫)	[1]	1	1	1	1	1	1	1	1	1	9	Yes
DMZ	Dòumén (Town) 斗門 (斗門鎮)	[1]	1	1	1	1	1	1	0	1	1	8	Yes
JM	Jiāngmén (Báishā) 江門 (白沙)	[1]	1	1	1	1	1	1	0	1	1	8	Yes
XH	Xīnhuì (City) 新會 (會城)	[1]	1	0	1	1	1	1	0	1	1	7	Yes
TS	Táishān (City) 台山 (台城)	[1]	1	0	1	1	1	1	0	1	1	7	Yes
KP	Kāipíng (Chìkǎn) 開平 (赤坎)	[1]	1	0	1	1	1	1	0	1	1	7	Yes
EP	Ēnpíng (Niújiāng) 恩平 (牛江)	[1]	1	0	1	0	1	1	0	1	0	5	Yes
HS	Hèshān (Yǎyáo) 鶴山 (雅瑤)	[1]	1	0	1	1	1	1	0	1	1	7	Yes
DG	Dōngguǎn (City) 東莞 (莞城)	[1]	1	1	1	1	1	1	1	1	1	9	Yes

Table 2.1: Yue Dialect Family Classification

Abbr	Full Name	Src	1	2	3	4	5	6	7	8	9	Sum	Yue?
BA	Bǎo'ān (Shājǐng) 寶安 (沙井)	[1]	0	1	1	1	1	1	1	1	1	8	Yes
HZ_h	Huìzhōu (City) 惠州 (市區)	[1]	0	0	1	1	1	1	0	1	0	5	Yes
QX_h	Dōngguǎn (Qīngxī) 東莞 (清溪)	[1]	0	0	1	0	1	0	0	1	0	3	No
STJ_h	Shēnzhèn (Shātóujiǎo) 深圳 (沙頭角)	[1]	0	0	1	0	0	0	0	0	0	1	No
LT_h	Cónghuà (Lǚtián) 從化 (呂田)	[1]	0	0	1	0	1	0	0	0	0	2	No
NL_h	Zhōngshān (Nánláng) 中山 (南蓢)	[1]	0	0	1	1	1	0	1	0	0	4	No
LD_m	Zhōngshān (Lóngdū) 中山 (隆都)	[1]	1	1	0	1	0	0	0	0	0	3	No
QY	Qīngyuǎn 清遠	[2]	1	1	1	1	1	1	1	1	1	9	Yes
FG	Fógāng 佛岡	[2]	1	1	1	1	1	1	1	1	1	9	Yes
YD	Yīngdé (Hánguāng) 英德 (浛洸)	[2]	1	1	1	1	1	1	1	1	1	9	Yes
YS	Yángshān 陽山	[2]	1	1	1	1	1	1	0	1	1	8	Yes
BT	Liánshān (Bùtián) 連山 (布田)	[2]	1	1	1	0	1	0	0	0	1	5	Yes
LX	Liánxiàn 連縣	[2]	1	1	1	0	0	1	0	1	0	5	Yes
SG	Sháoguān 韶關	[2]	1	1	1	1	1	1	1	1	1	9	Yes
QJ	Qǔjiāng (Mǎbà) 曲江 (馬壩)	[2]	1	1	1	1	1	1	1	1	1	9	Yes
RH	Rénhuà 仁化	[2]	1	1	1	1	1	1	1	1	1	9	Yes
LC	Lèchāng 樂昌	[2]	1	1	1	1	1	1	1	1	0	8	Yes
ZQ	Zhàoqìng (Gāoyào) 肇慶 (高要)	[3]	0	1	1	1	1	1	1	1	1	8	Yes
SH	Sìhuì 四會	[3]	1	1	1	1	1	1	1	1	1	9	Yes
GN	Guǎngníng 廣寧	[3]	1	1	1	1	1	1	0	1	0	7	Yes
DQ	Déqìng 德慶	[3]	1	1	1	0	1	1	1	1	1	8	Yes
HJ	Huáijí 懷集	[3]	1	1	1	1	1	1	1	1	1	9	Yes

Table 2.1: Yue Dialect Family Classification

Abbr	Full Name	Src	1	2	3	4	5	6	7	8	9	Sum	Yue?
FK	Fēngkāi (Nánfēng) 封開 (南豐)	[3]	1	1	1	1	1	1	0	0	1	7	Yes
YF	Yúnfú (City) 雲浮 (雲城)	[3]	1	1	1	1	1	1	1	1	1	9	Yes
XX	Xīnxīng 新興	[3]	1	1	1	0	1	1	0	1	1	7	Yes
LD	Luódìng 羅定	[3]	1	1	1	1	1	1	1	1	0	8	Yes
YN	Yùnán (Píngtái) 郁南 (平台)	[3]	1	1	1	0	1	1	1	1	1	8	Yes
XY	Xìnyí 信宜	[4]	1	1	1	1	0	1	1	0	0	6	Yes
LJ	Liánjiāng 廉江	[4]	1	1	1	1	1	1	1	1	1	9	Yes
DC	Táishān Dàncūn 台山淡村	[5]	1	0	1	1	1	1	0	1	1	7	Yes
GP	Guìpíng (Cty. Seat) 桂平縣城	[6]	1	1	1	1	1	1	1	1	1	9	Yes
MS	Méngshān (Xīhé) 蒙山西河	[6]	0	1	0	1	1	1	1	0	1	6	Yes
WX_ng	Wǔxuān (Zhāngcūn) 武宣樟村	[6]	0	1	0	0	0	0	1	0	0	2	No
CZ_p	Chóngzuǒ (Sìpái) 崇左四排	[6]	1	0	1	1	1	0	1	0	0	5	Yes
GG	Guìgǎng (Nánjiāng) 貴港南江	[6]	0	1	1	1	1	1	1	0	1	7	Yes
BL	Běiliú (Tángliáo) 北流唐僚	[6]	1	1	1	1	1	1	1	0	1	8	Yes
BS	Bósè (City) 百色市內	[6]	1	1	1	1	1	1	1	1	1	9	Yes
BB	Bóbái (Cty. Seat) 博白縣城	[6]	1	1	1	1	1	0	0	1	1	7	Yes
LS	Língshān (Cty. Seat) 靈山縣城	[6]	0	1	1	1	1	0	1	0	1	6	Yes
PB	Pǔběi (Cty. Seat) 浦北縣城	[6]	1	1	1	0	0	0	1	1	1	6	Yes
QZ	Qīnzhōu (City) 欽州市內	[6]	1	1	1	1	1	1	1	0	1	8	Yes
BH	Běihǎi (City) 北海市內	[6]	1	1	1	1	0	1	1	1	1	8	Yes
NM	Níngmíng	[6]	1	1	1	1	1	1	1	1	1	9	Yes
NN	Nánníng (City) 南寧市內	[6]	1	1	1	1	1	1	1	1	1	9	Yes

Table 2.1: Yue Dialect Family Classification

Abbr	Full Name	Src	1	2	3	4	5	6	7	8	9	Sum	Yue?
NN_p	Nánníng (Shājǐng) 南寧沙井	[6]	1	1	1	1	1	0	1	1	1	8	Yes
FS_p	Fúsuí (Lóngtóu) 扶綏龍頭	[6]	1	1	1	0	1	0	0	1	1	6	Yes
BS_p	Bósè (Nàbì) 百色那畢	[6]	1	1	1	1	1	0	1	1	1	8	Yes
BY	Bīnyáng (Xīnqiáo) 賓陽新橋	[6]	1	1	1	1	0	0	1	1	0	6	Yes
HX	Héngxiàn (Cty. Seat) 橫縣縣城	[6]	1	1	1	0	0	0	1	0	1	5	Yes
BL_h	Běiliú (Táng'àn) 北流塘岸	[6]	0	1	1	0	0	0	0	0	0	2	No
MS_h	Mǎshān (Piānlián) 馬山片聯	[6]	0	1	0	0	0	0	1	0	0	2	No
PN_m	Píngnán (Yǎbù) 平南 (雅埠)	[6]	0	0	0	0	0	0	0	0	0	0	No
LG_g	Línguì 臨桂	[6]	0	0	0	0	0	0	0	0	0	0	No
LR_g	Luòróng 雒容	[6]	0	0	1	0	0	0	0	0	0	1	No
GLCY_p	Guìlín (Cháoyáng) 桂林朝陽	[6]	0	0	0	0	0	0	1	0	0	1	No
LC_p	Língchuān (Tánxià) 靈川潭下	[6]	0	0	0	0	0	0	0	0	0	0	No
YF_p	Yǒngfú (Táochéng) 永福桃城	[6]	1	1	0	0	0	1	0	0	0	3	No
GLYS_p	Guìlín (Yànshān) 桂林雁山	[6]	0	0	0	0	0	0	0	0	0	0	No
YS_p	Yángshuò (Jìmǎ) 陽朔驥馬	[6]	0	1	0	0	0	0	0	0	0	1	No
FC7_t	Fùchuān (Qīdù) 富川七都	[6]	0	1	0	0	0	0	0	0	0	1	No
FC8_t	Fùchuān (Bādù) 富川八都	[6]	0	1	0	0	0	0	0	0	0	1	No
HZ_t	Hèzhōu (Jiǔdū) 賀州九都	[6]	0	0	1	0	0	0	0	0	0	1	No
LS_p	Lóngshèng (Hóngyáo) 龍勝紅瑤	[6]	0	1	0	0	0	0	0	0	0	1	No
DX_t	"Dàoxiàn Húnán 湖南道縣"	[6]	0	1	0	0	0	0	0	0	0	1	No
XD	Xìndū (Pūmén) 信都鋪門	[6]	1	1	1	1	1	0	0	0	0	5	Yes
ZS_t	Zhōngshān (Gōng'ān) 鐘山公安	[6]	0	1	0	0	1	0	1	0	0	3	No

Table 2.1: Yue Dialect Family Classification

Abbr	Full Name	Src	1	2	3	4	5	6	7	8	9	Sum	Yue?
PL_t	Pínglè (Zhāngjiā) 平樂張家	[6]	1	0	0	0	1	0	0	0	0	2	No
WT_t	Línguì (Wǔtōng) 臨桂五通	[6]	1	1	0	0	0	0	0	0	0	2	No
RS_t	Róngshuǐ (Cty. Seat) 融水縣城	[6]	0	1	1	0	1	1	0	0	0	4	No
YZ_b	Yízhōu (Déshèng) 宜州德勝	[6]	1	1	0	0	0	0	0	0	0	2	No
GY_x	Guānyáng (Wénshì) 灌陽文市	[6]	0	0	0	0	0	0	0	0	0	0	No
QZ_x	Quánzhōu (Cty. Seat) 全州縣城	[6]	0	0	0	0	0	0	0	0	0	0	No
WC	Wúchuān 吳川	[7]	1	1	1	1	1	0	0	1	1	7	Yes
HZ	Huàzhōu (Downstream) 化州下江	[7]	1	1	1	1	1	0	0	1	1	7	Yes
YJ	Yángjiāng 陽江	[8]	1	1	1	1	1	1	0	1	1	8	Yes

There is one obvious false positive in this data. The Huìzhōu dialect recorded in Zhān, Cheung, et al. (1987a) is classified as a Hakka dialect without controversy. There may be other false positives that are difficult to detect because there is less consensus on the ultimate classification, as is the case for the Northern Guǎngdōng Patois and Pínghuà dialects. Interestingly, though, the Patois and Northern Pínghuà dialects score well below the threshold to classify them as part of the Yuè family. The Southern Pínghuà dialects score higher on the whole. This supports the widely held view that Pínghuà does not constitute a valid classification.

The greatest number of uncertain dialects are located in the north near the border with Húnán. As it turns out, in our internal grouping, a good number of these dialects

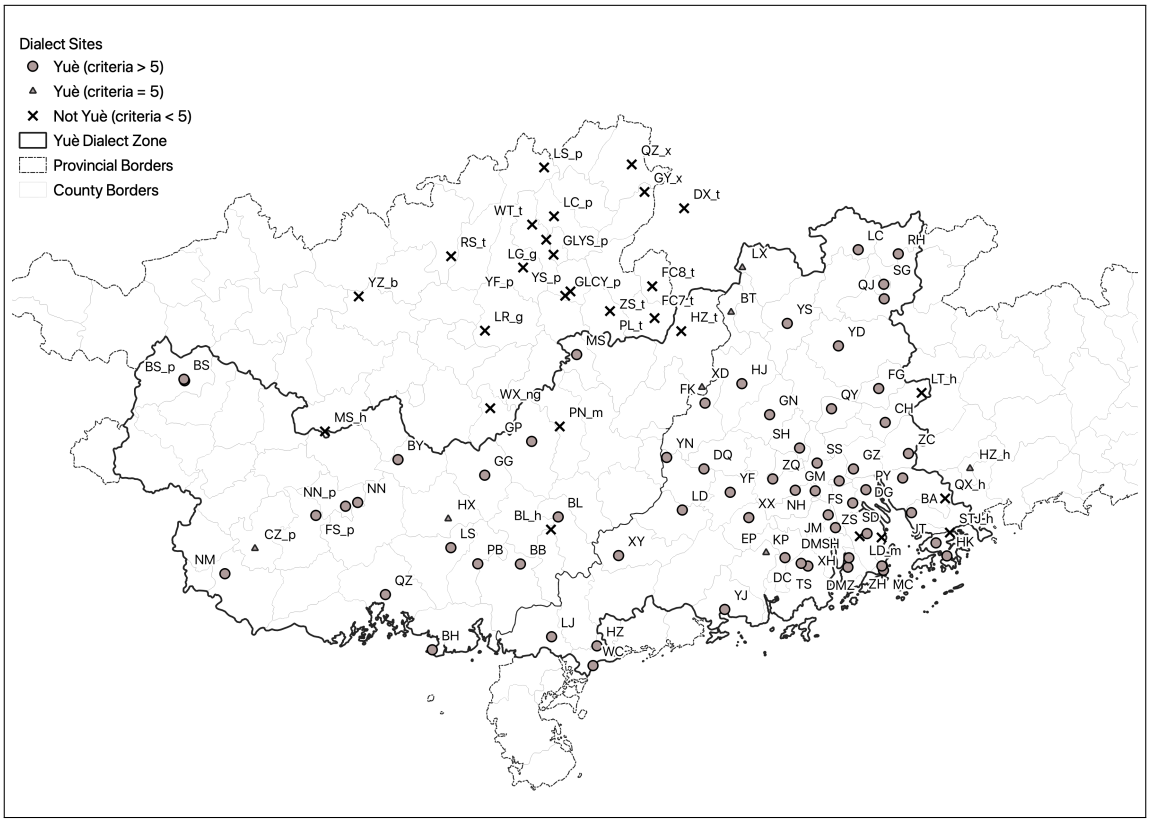


Figure 2.1: The Geographic Distribution of the Yuè Dialects

are linguistic islands and likely are the product of convergence with other dialect groups in the area. For the nonce, we will accept them as part of the Yuè family nonetheless.

2.2 Internal Grouping

While the classification of dialects into groups is not a prerequisite to applying the comparative method, it is nonetheless useful in our case. 70 sites are too unwieldy to present as correspondences. Moreover, many of the sites in our data are so phonologically similar that comparison between them is trivial and only overshadows more important corre-

spondences. It would be better to present a smaller sample of dialects that still adequately represent the diversity of our chosen corpus. We can accomplish this by taking samples of dialects in each group of the family, thereby guaranteeing diversity. To derive a classification, we use computational methods in dialectometry to measure the lexical distances between all 70 sites and then cluster the sites into groups according to those distances.

Dialectometry is a branch of the field of dialectology that seeks to measure linguistic differences. Dialectometry was born when Jean Seguy (1971) first attempted to quantify linguistic distance by counting overlapping responses in dialect vocabulary surveys. Hans Goebl (1982, 1984) seized upon Seguy's breakthrough, making many refinements, such as accounting for the relative frequency of words in surveys. It is from this point forward that dialectometry began to garner broad scholarly attention.

2.2.1 Measuring the Linguistic Distances

In modern dialectometry, it is possible to measure linguistic distance in terms of both phonetic and lexical distance. We opt to measure the latter. To measure phonetic distance, one has to assemble pronunciations of common words across all dialects. Usually, this sort of data can be gathered when producing texts in fieldwork surveys of dialect areas. Words that remain common in all sites can be compared phonetically. Unfortunately, the corpus used for Common Yuè does not include texts for most of the sites. Another way of collecting similar pronunciations in the case of Chinese might be comparing character reading data. However, we have decided not to take this approach. Often times, char-

acter readings do not represent actual words. This is particularly true in the fieldwork surveys produced by Zhān and Cheung at al.'s teams. Often times, the pronunciation of characters inferred by the way they are pronounced in the context of actual words completely differs from what is listed as character readings in their work. Rather than hunt for common words in the documentation of the lexicon, many of which seem to be words of recent origin, we have elected to measure lexical distance instead.

Yuè dialects present a unique problem for measuring lexical distance. Conventionally, lexical distance is measured by counting identical responses to surveys written in a standard orthography. Yuè dialects do not have a standard orthography. Although the fieldworkers behind the studies attempted to write elicited lexical forms in Chinese characters, the conventions they used vary widely. For example, the word for "son," which takes the general shape /tsɐi^{3}/, is written as both 仔 and 崽 by different fieldworkers. This makes the surveys' orthography incomparable. Some words that have no clear sinitic cognate are not written with characters at all, like the word for "viscous," which takes the general form /kit^{8}/. The only objective way to measure the lexical documentation, then, is by comparing phonetic transcriptions. In their study of lexical distance of American English dialects as recorded in the *Linguistic Atlas of the Middle and South Atlantic States* (LAMSAS), Nerbonne and Kleiweg (2003) used Levenshtein Distance to measure the edit distance of strings of characters used to write survey responses. They did this to approximate lemmatization in cases where the answers were clearly related forms of words like "fair off" versus "fairs off." We will do the same with the phonetic transcriptions of the

words in the surveys. We have used Peter Kleiweg's dialectometry and cartography software package, RuG/L04, to make these measurements as well as conduct further analysis and produce maps[1].

Levenshtein Distance

Levenshtein Distance is a sequence comparison algorithm that measures the relative disparity, or distance, of sequences (Levenshtein, 1966). It seeks the least costly set of basic operations necessary to transform one sequence into another. The basic operations are element insertion, deletion, and substitution. Each operation is given a "cost", or "weight." In a basic implementation of the algorithm, like the one we use, all costs are set to one. The cost of the entire transformation is the sum of the costs of all the operations involved in the transformation.

In the case of measuring the distance between two strings of characters, the elements of the sequences are the characters themselves. In the case of phonetic transcription, there are many ways to choose what constitutes an element in a sequence. It can be the literal phonetic symbols, phonetic segments, or series of distinctive features. To make a simple illustration of Levenshtein Distance, we will perform the operations on just the phonetic symbols themselves, with diacritics and digraphs counted as separate elements. Consider the word for "thief" in the Bósè 百色 (BS) dialect, [tʃʰɐk^{22}], and the Dàncūn 淡村 dialect (DC) , [tʰak^{32}]. The minimal set operations to transform the string of phonetic

[1]RuG/L04 is available for download at `http://www.let.rug.nl/~kleiweg/L04/` along with extensive documentation and tutorials. I am indebted to Prof. Kleiweg's generosity.

symbols that writes the Bósè word in to the Dàncūn word is demonstrated below.

t	ʃ	ʰ	ɐ	k	²	²		
t		ʰ	ɐ	k	²	²	delete [ʃ]	1
t		ʰ	a	k	²	²	substitute [a]	1
t		ʰ	a	k	³	²	substitute [³]	1
								3

Because transforming sequences of different lengths involves varying costs, it is necessary to normalize Levenshtein Distance when comparing distances between sequences. In this study, we use the following formula to normalize the distances, where f is the cost of transformation, l_1 is the length of the first sequence and l_2 is the length of the second sequence.

$$\frac{f}{l_1 + l_2} \tag{2.1}$$

The advantage of the Levenshtein Distance in dialectometry is that it allows for numerical treatment of data rather than just a categorical treatment Nerbonne and Kretzschmar (2003). This is particularly useful in our case with Yuè dialects since the categorical lemmatization judgements reflected in the Chinese characters investigations chose to write their data in is not comparable across studies. Application of the Levenshtein Distance to measure the distance between phonetic transcriptions has proven reliable and fruitful many times over (Nerbonne, Heeringa, et al., 1996; Nerbonne and Heeringa, 1997; Heeringa, 2004). Its use as an appropriate measure of lexical similarity was demonstrated in Nerbonne and Kleiweg (2003)'s study of LAMSAS. It is an appropriate way to measure

the Yuè data as well, but does have its limitations. For example, compounds in which the elements are inverted, like /kɐw³ kuŋ¹/ 狗公 versus /kuŋ¹ kɐw³/ 公狗 cannot be readily uninverted and aligned in an economical way due to the basic nature of the transformation operations. The Levenshtein Distance algorithm overestimates these sorts of sequences' relative distances. Fortunately, this type of inversion is rare in our data.

Heeringa (2004) demonstrated the feasibility of measuring the edits of phonological features rather than measuring transcriptions themselves directly. In this study, we have ultimately opted for a similar approach where we tokenize phonetic segments and encode them in terms of distinctive features. The features chosen are not an exhaustive description, but rather a minimal functional description that can measure phonetic distance well enough to judge lexical similarity. A full listing of the feature encoding of the various segments is included in Appendix A.

Fieldworker Bias

Sometimes, differences in the technical aspects of fieldwork, like elicitation and transcription, can distort the linguistic differences measured in dialectometry. This sort of bias can create "fieldworker isoglosses." Often times, dialectologists will try to control for variables in fieldwork as much as possible by limiting their sources to just one fieldworker. We do not have this luxury when it comes to Yuè dialects. The investigators in most of the surveys were not transparent about their distribution of the work and other details of the investigation itself.

To account for differences in transcription practices, Wieling and Nerbonne (2011) proposed a procedure by which phonetic symbols are iteratively merged according to pointwise mutual information (PMI) scores of alignments generated by the Levenshtein algorithm. Different phonetic symbols that align with each other more than chance are merged iteratively according to which pairs are more strongly likely to occur together in an alignment. This method helps reconcile fieldwork in which investigators used different sets of phonetic symbols.

Unfortunately, Wieling and Nerbonne's procedure does not help eliminate transcription bias in our data. Between the Yuè dialect surveys, the greatest difference is not the set of phonetic symbols used itself but the phonemic analysis of the dialects. For example, the Zhān and Cheung et al. surveys analyze what they describe phonetically as [ɪŋ] and [ɪk] phonemically as /eŋ/ and /ek/, while Xiè (2007) analyzes these same finals as /iŋ/ and /ik/. Similarly, in the same studies, alveolar and post-alveolar places of articulation for initial consonants are in complementary distribution in the vast majority of the dialect sites. The Zhān and Cheung surveys write these with phonemic alveolar sibilants, but Xiè does not group them as one phoneme despite them being in complementary distribution and writes them separately as alveolar and post-alveolar sibilants. These differences in transcription practice create artificial distances. We opted to "sanitize" the transcriptions by making the following changes before comparison:

1. Rewrite <e> as <i> and <o> as <u> before velar codas.

2. Rewrite <ɕ>, <ʃ>, and <ʂ> as <s>

3. Rewrite cases of initial j plus medial i and zero initial plus medial i as j

4. Rewrite cases of initial w plus medial u and zero initial plus medial u as w

5. Rewrite cases of <niɐ> as <ȵɐ>

In an extreme minority of cases, these rewrites result in phonological mergers. However, in most cases, they do not.

Beyond just differences in transcription, there is likely also some degree of fieldworker bias present in elicitation. The Zhān and Cheung et al. surveys are fairly consistent in the number of dialectically appropriate responses they received for each word they attempted to elicit. In most cases, it is just one response, but occasionally there are two or even three. Xiè (2007), on the other hand, despite attempting to elicit the same words as Zhān, Cheung, et al. (1987a) almost always gives just one response for each word. This suggests that their methods of elicitation differed. Unfortunately, there is no way to normalize this effect. The smaller studies we reference did not always elicit all the words the larger surveys did. Fortunately, this is not too much of a problem. Finally, the average distance of all responses for each pair of sites is taken. If a few responses are missing, it does not affect the overall estimated distance too greatly.

Results

We measured the lexical distance of all pairs of the 70 Yuè dialect sites we identified in the previous study. To obtain the average lexical distance between all pairs, we took the average of the normalized Levenshtein Distance of the feature-encoded tokenized strings of the original transcriptions of the responses in the lexical survey. We chose a total of 77 words to sample. Many of these words occur in the lexical comparison maps in the middle of Zhān et al. (2000, pp.225–290). Others we chose arbitrarily. The sample is intended to reflect general lexical differences and is not restricted to any one lexical domain. A listing of all lexical items and their index numbers in the surveys is given in Appendix B.

When comparing words for which there were multiple responses, we employ Nerbonne and Kleiweg (2003, p.349)'s method to compare all forms. For each response for each site, we find a pair with a response from the opposing site that is of minimal distance. Then we take the average of this set of pairs.

To make sure our chosen method was the most accurate, we compared it to other methods. Nerbonne and Kleiweg (2007) proposed a metric of the quality of measurement of linguistic differences. They call this metric Local Incoherence and argue that it is a good "yardstick" for dialectometry because measures of higher quality will minimize the linguistic incoherence, or linguistic distance, between proximate locations. We took several measurements of distance with different methods and used Nerbonne and Kleiweg's equation for Local Incoherence to compare the quality of each.

Calculating the lexical distance between dialects using Levenshtein Distance applied

Raw String	4.230
Sanitized String	4.195
Raw String Tokenized and Encoded for Features	4.208
Sanitized String Tokenized and Encoded for Features	**4.193**

Table 2.2: Local Incoherence of different metrics

to sanitized tokenized segments encoded by distinctive features proved to best minimize local incoherence, so it is the distances calculated using this method that we used in the internal grouping analysis. The final result of our calculation is a distance matrix of the average lexical distance between all pairs of dialect sites.

2.2.2 Further Analysis

Now that we have obtained a distance matrix for all of the sites, we can engage in more sophisticated types of analysis to produce an internal grouping as well as create geographic visualizations of the dialect boundaries.

Multidimensional Scaling

Multidimensional scaling (MDS) is a technique that reconstructs relative positions of points in an arbitrary number of dimensions based on a square matrix of the distances between each point. Scholars of dialectometry have successfully used this technique to visualize the relative linguistic distance between dialects (Nerbonne and Kleiweg, 2003; Heeringa, 2004; Valls et al., 2012; Jeszenszky et al., 2019). Normally, this is done in either two or three dimensions. Two dimensions can be represented in a scatter plot along

an X and Y axis. Three dimensions are conventionally mapped to color values, such as RGB (red, blue, green) and then projected on to a geographic map.

While MDS alone does not divide dialects into discrete groups, it represents dialectal continua well. As dialects gradually vary along a continuum, colors change in hue. This is a powerful visual aide, but still suffers some problems. In three dimensions, it is best suited to represent up to eight different groups as extremes of the color cube. The colors derived from the MDS analysis can also pose unnatural contrasts for the human eye. Changing which axis represents which color can have a dramatic impact despite representing the same data. These issues can make it especially difficult for colorblind people to see the relevant dialect boundaries (Kleiweg, 2007b).

We convert the dimension vectors of our MDS into colors by scaling each dimension to a range of 0 to 1. We interpret these values in the CIE RGB colorspace, which was intended to reflect the perceivable range of color in human vision (Smith and Guild, 1931). These color values are more natural to the human eye than a direct mapping to RGB, but might still pose some challenges for colorblind people.

The variety of MDS we use here is Kruskal's Non-Metric Multidimensional Scaling (Kruskal, 1964). We use a Principal Components Analysis as the initial embedding, i.e. the first estimate of relative positions of points, and then run 50 iterations of the method updating the embedding each time until convergence is achieved. Our final r value was 0.9036, indicating that the final embedding accounts for more than 90% of the variance present in the data.

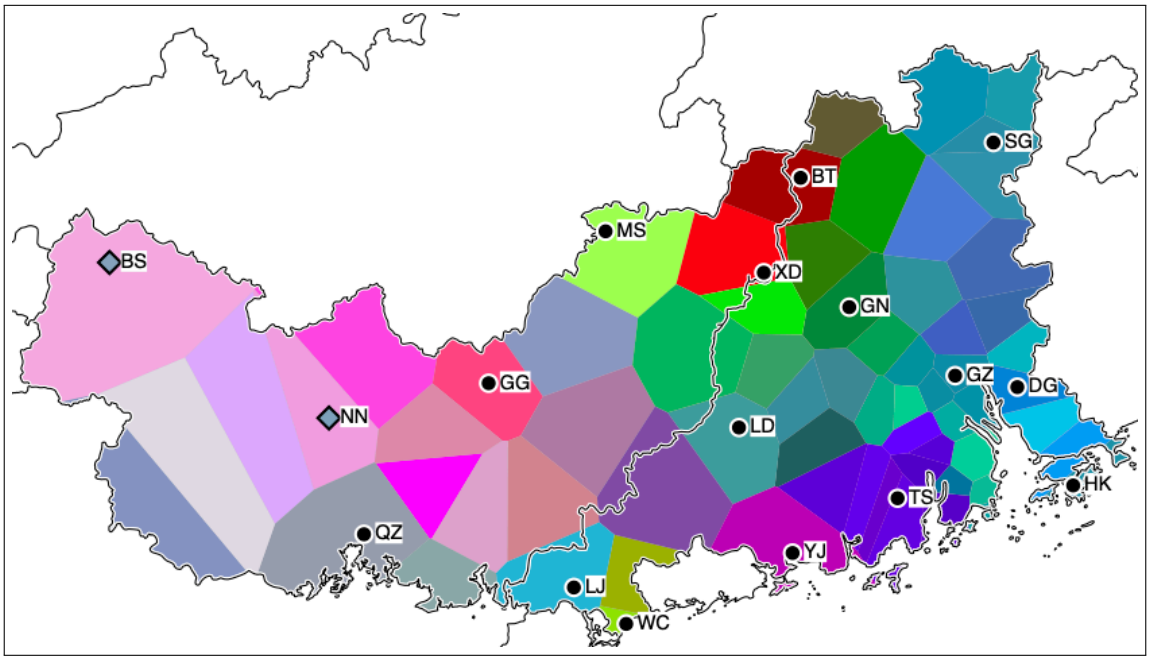


Figure 2.2: Lexical distance between Yuè dialect site represented by color

Figure 2.2 is the final result of this analysis. The dialects seem to naturally group into Eastern, more blue-green, and Western, more pink, groups. The Wǔyì dialects form their own unique group, represented in deep blue, as Yue-Hashimoto proposed. However, the Yángjiāng dialect does not seem to bear as close as a relationship to the Wǔyì group as she suggested. The Liángjiāng dialect is close in color to Dōngguǎn, also as she proposed, but it is difficult to discern if those should be grouped any closer than any of the similarly colored dialects in the east. Interestingly, in the north, there appear to be a few islands, Méngshān, Xìndū, and Bùtián, that do not appear to bear a relationship with the surrounding dialects nor each other. Finally, in the West it is important to note that the Bósè and Nánníng urban dialects differ from their rural "Pínghuà" counterparts and appear to

be closer to the Níngmíng dialect. I have marked them as diamonds in their respective locations.

Hierarchical and Clustering

Hierarchical Clustering is a technique that seeks to classify elements of a set into meaningful nested subsets. The result is a tree, similar to the trees historical linguists draw to visualize the divergence of language families. Because of the similarity, it warrants saying here that the hierarchies of Hierarchical Clustering based on lexical distance do not necessarily reflect historical developments. When we calculate how close two dialects are together, we do not take special considerations to only count common innovations and exclude common retentions. Similarly, we do not attempt cases of common innovation from later contact and diffusion. The hierarchies in our cluster then represent how far away certain dialects are from others in a given cluster.

Clustering algorithms involve placing the points of the shortest relative distance to each other in common groups. In hierarchical clustering, points and groups themselves can also be placed in other groups. Finding the distance between clusters can be calculated in many ways, but the way it is typically done in dialectometry is finding the distance between the mean position of two clusters (Heeringa, 2004; Nerbonne, Kleiweg, et al., 2008). This leaves us with two clustering algorithms to work with: Unweighted Pair Group Method using Arithmetic averages (UPGMA) and Weighted Pair Group Method using Arithmetic averages (WPGMA). The difference between these algorithms is that the

unweighted variety accounts for the relative size of clusters when combining them, while the weighted variety treats all clusters as if they were equal weight (Heeringa, 2004, p.150). The choice of which algorithm to use is a matter of which best reflects relevant distinctions in the data. We will revisit this issue shortly.

Ward's Method can be a useful diagnostic tool for working out the best clustering (Kleiweg et al., 2004; Kleiweg, 2007a). Ward's method decides which points to group together by finding the minimum variance between points and a calculated average point of a group. It tends to generate clusters of uniform size that can be unnatural, so it is only a starting point rather than the final grouping itself (Heeringa, 2004, p.150).

Figure 2.3 is a dendrogram representing Ward's Method applied to the Yuè dialect distance matrix. Figure 2.4 demonstrates how some of the groups derived by Ward's Method are arbitrary. Outlier dialects, like MS, XD, YJ, BT, FK, etc. cluster with larger groups despite being too distant to cluster with them naturally. Excluding these outliers, the groups appear to mostly have significant space between them in 2.4a. The cyan group of Northern sites looks suspect, but this is because of the stress of projecting the dialect distances in only two dimensions. Taking out the WC and HZ dialects removes some of this stress. There is a more or less clear divide in 2.4b and the West-East split (yellow and blue from the rest) is more apparent. Focusing on just the Eastern sites in 2.4c, it is clear the red and green groups ought to be separate. The red and cyan groups also appear to have a significant distinction, but it is not clear if the place where Ward's Method drew the boundary is correct or not. XX, for example, seems far enough between the two

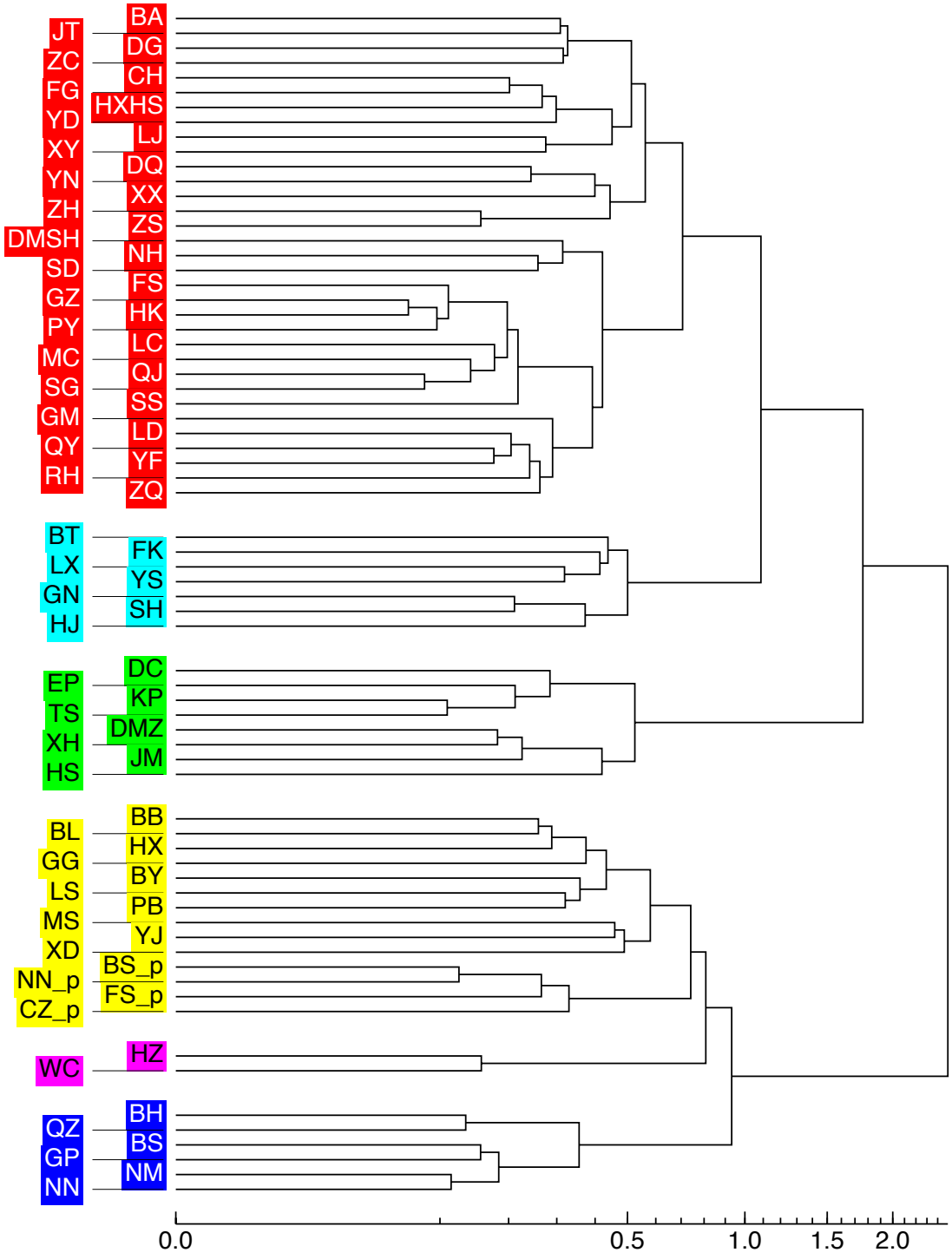
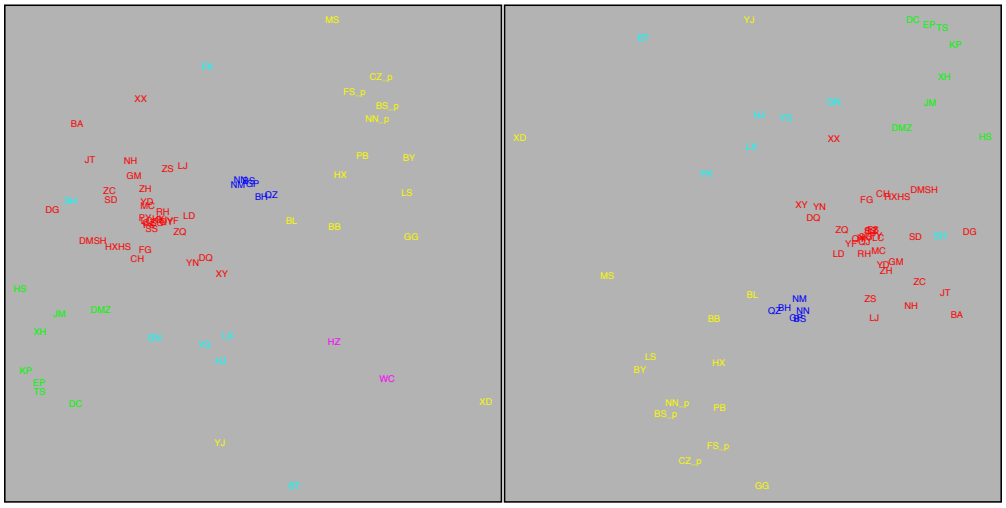
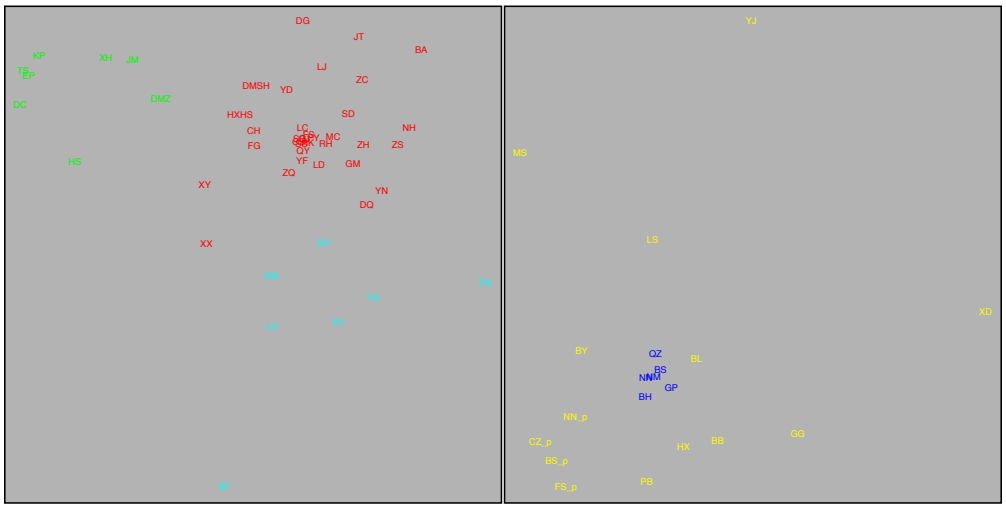


Figure 2.3: The 70 Yuè dialect sites clustered into six groups according to Ward's Method



(a) All Sites

(b) Excluding HZ, WC



(c) Eastern Sites

(d) Western Sites

Figure 2.4: MDS 2d Projection of Ward's Method Groups

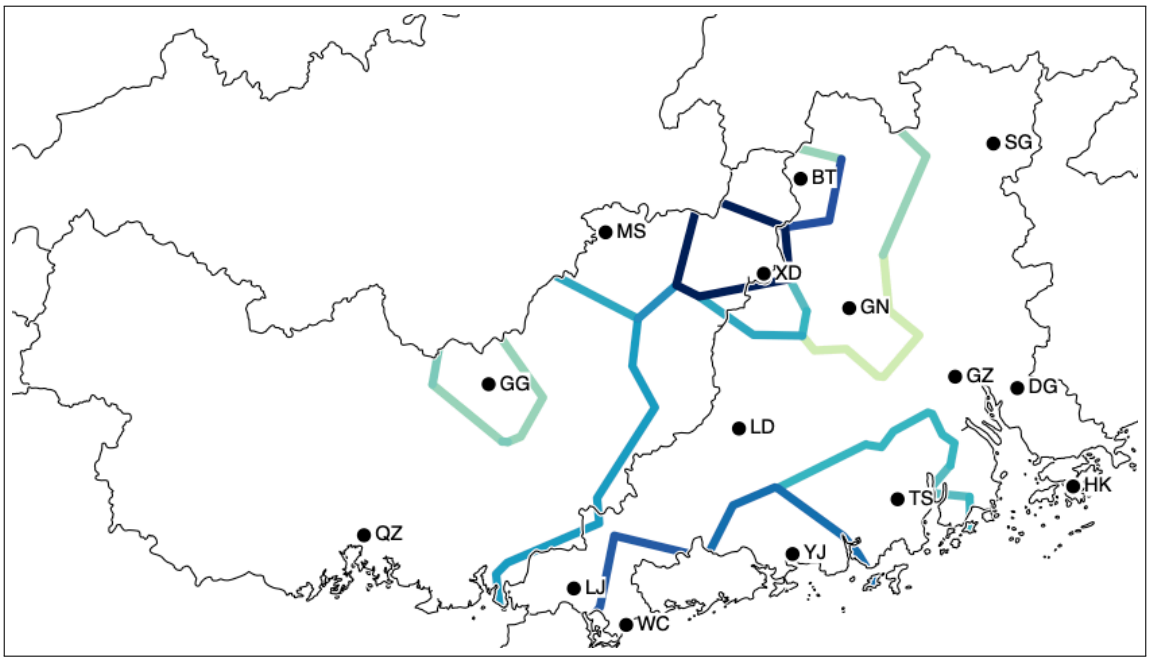
groups that it could go either way. Finally, in 2.4d, it becomes clear that Ward's Method's subdivision of the Western sites is arbitrary. The blue group is not significantly separate from the blue group. These insights will help us choose between the UPGMA and WPGMA clustering methods.

Composite Cluster Maps

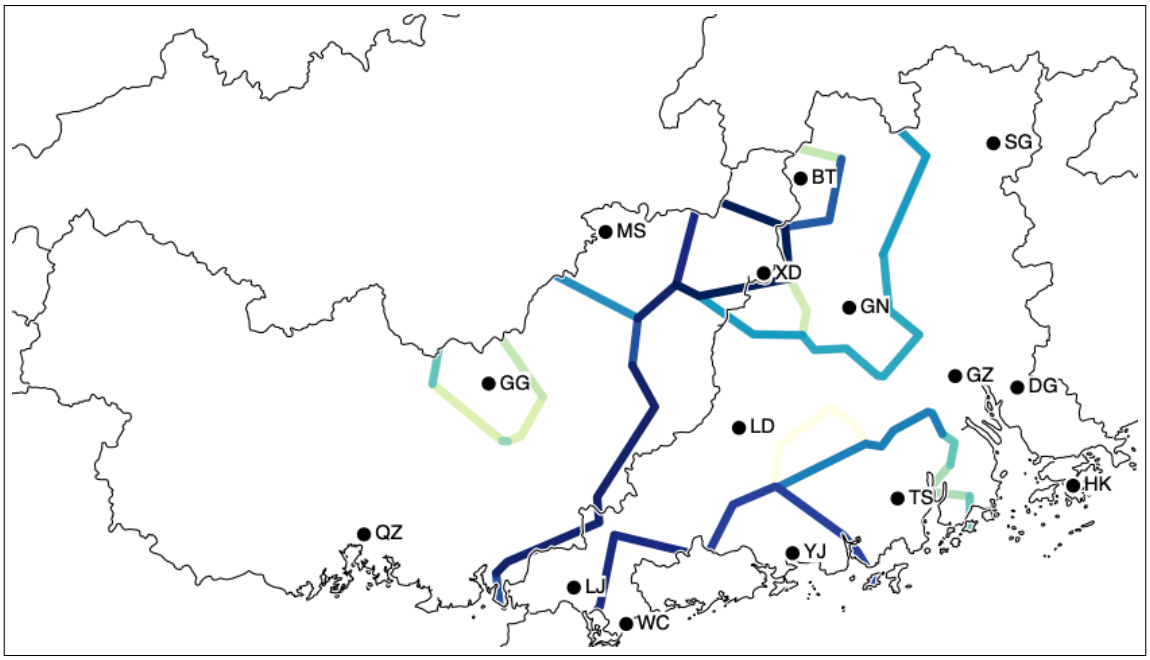
Simple hierarchical clustering, like the kind in Figure 2.3, can be unstable. In the case of large dialect surveys, small fluctuations in distance matrices can drastically change the structure of the clusters. To overcome this problem, we can build composite clusters by adding noise to the distance matrix and combining the results of many clustering runs (Kleiweg et al., 2004; Nerbonne, Kleiweg, et al., 2008). This technique takes advantage of the inherent instability of hierarchical clustering algorithms and our data to highlight naturally strong borders along which dialects sites that show up consistently despite adding noise to the distance matrix (Kleiweg et al., 2004).

We can then project these borders geographically using a special type of map that Kleiweg dubs a "cluster composite map." This map draws borders between dialect sites of varying shade depending on the cophenetic distance between the sites resulting from the composite of "noisy clustering." The darker the border, the more distant the sites are in the composite cluster.

Figure 2.5 shows two composite cluster maps using both UPGMA and WPGMA. The borders they draw are nearly identical. The main difference is the weight of the border



(a) UPGMA



(b) WPGMA

Figure 2.5: Composite Cluster Maps of the Yuè dialect sites. Clusters: 12, Noise: 1, Repetitions: 100, Contrast: 4

that separates the northern group relative to the northern island sites XD and MS and the western group. Both methods place XD on a branch very distantly separated from all sites. This is reflected to an extent in our MDS plots and color map. However, in the UPGMA map, FK is more distant from the northern group than the northern group is from the eastern group. This does not appear to be true in 2.4b. GG in the UPGMA map also appears as distant from the western group as the northern group is from the eastern group. Although there is considerable distance between GG and the other western sites in 2.4, it does not seem to be as great as the relative separation of the northern group from the eastern group. The WPGMA map reflects the situation we explored using the MDS plots and map more accurately.

Finally, we can use the cophenetic distances from composite clustering as a distance matrix itself to perform MDS (Kleiweg, 2007c). This allows us to visual the groups derived from composite clustering while still partially seeing the continua within each group. In 2.6 we use the Kleiweg (2007c)'s recommended parameters for such a map and combine the UPGMA and WPGMA composite clusters cophenetic distances.

2.2.3 Conclusion

Based on the above analysis, we group the Yuè dialects into an eastern and western region, then subdivide the eastern region into three groups. The western region forms its own group. The groups are as follows.

- Eastern (粵東):

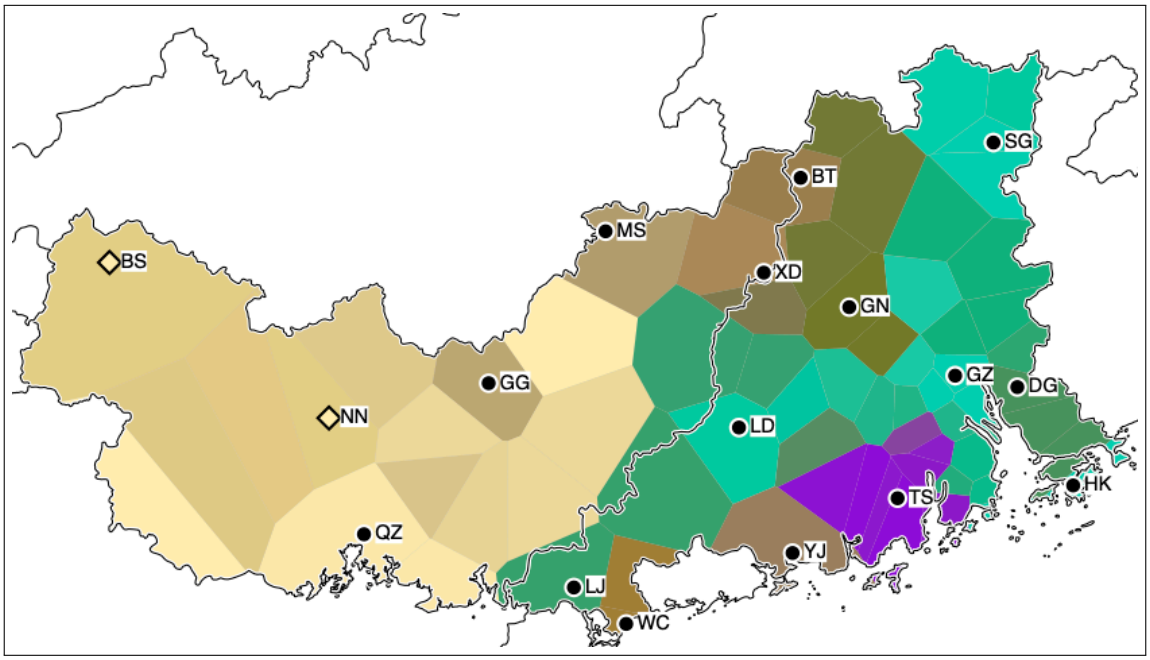


Figure 2.6: A 3d MDS projection based on the cophenetic distances of composite clustering

1. Fēnglián Group (封連片). Distributed just south of the Nánlǐng 南嶺 mountains in inland western Guǎngdōng and eastern Guǎngxī, extending as far as Sìhuì 四會 (SH). Named for the old prefectures Liánzhōu 連州 and Fēngzhōu 封州 that occupy the areas around modern Liánxiàn 連縣 and Fēngkāi 封開. (Brown Green in Fig. 2.6).

2. Guǎngfǔ Group (廣府片). Distributed in the Pearl River Delta and along the Western and Northern (Běijiāng 北江) Rivers. Named for administrative center in Guǎngzhōu. (Light Green in Fig. 2.6).

3. Sìyì Group (四邑片). Distributed around the cities of Táishān 台山, Kāipíng, 開平 Xīnhuì 新會, Ēnpíng 恩平, and Hèshān 鶴山. Named for the endonym of

the cultural group in lieu of the modern administrative rebranding Wǔyì 五邑. (Purple in Fig. 2.6).

- Western (粵西):

4. Yōngxún Group (邕潯片). Distributed along the Yōng (邕江), Yù (鬱江) and Xún (潯江) rivers as well as the coastal western area around Qīnzhōu 欽州 and Běihǎi 北海. (Tan in Fig. 2.6).

The Yōngxún group might be further subdivided between a subgroup in the coastal area, Guìpíng, Níngmíng, and the Nánníng and Bósè city centers that is more Guǎngfǔ-like due to more recent waves of immigration (lighter in Fig. 2.6, and a subgroup that is more inland and isolated (darker in in Fig. 2.6). However, our analysis did not provide strong support for this level of subdivision.

Some dialects we leave unclassified for the time being. The documentation of the Yuè dialects is sparse outside of the Pearl River Delta region. There is not enough information to be certain how some of the more linguistically diverse outlying dialect sites tend to group.

It is clear that the Wúchuān and Huàzhōu dialects are closely related, but they form a sort of island and are not closely connected with other dialects in the region Lǐ (2014) mentions that these dialects have been influenced by Mín dialects and Zhuàng language and they are indeed sandwiched between predominately Mín speaking areas on the Léizhōu

(雷州) peninsula Diànbái (電白) district. Its linguistic distance may be the result of contact.

The same is likely true of the Yángjiāng dialect. It is an island in our data, but is reported to be similar to the neighboring Yángchūn dialect. As noted previously, it does not seem to share as close a connection with the Sìyì group as Yue-Hashimoto proposed.

There are a few dialects in the far inland north that also remain resistant to classification. The Xìndū, Méngshān, Bùtián, Fēngkāi, and Guìgǎng dialects all significantly lexically distant from the dialect groups that surround them. For the time being, we treat FK and BT as part of the Fēnglián group and GG as part of the Yōngxún group, but keep MS and XD separate as unclassified dialect sites. They both border the region of the Northern Patois dialects and probably also represent contact zones.

Table 2.3 lists all 70 dialect sites under our new classification scheme. The ones we use in our correspondence sets for Common Yue are written in bold and italics.

Region	Group	Sites
Eastern	Fēnglián	***YS***, BT, ***LX***, SH, GN, ***HJ***, ***FK***
	Guǎngfǔ	***GZ***, HK, JT, MC, PY, HXHS, CH, ZC, FS, ***NH***, SD, GM, ZS, ZH, DMSH, ***DG***, BA, QY, ***FG***, YD, SG, QJ, RH, LC, ZQ, SH, DQ, ***YF***, XX, LD, YN, XY, LJ
	Sìyì	JM, DMZ, ***XH***, ***TS***, ***KP***, ***EP***, HS, DC
Western	Yōngxún	GP, CZ_p, GG, ***BL***, BS, BB, LS, PB, ***QZ***, BH, NM, NN, ***NN_p***, FS_p, ***BS_p***, ***BY***, HX
Unclassified	Inland	XD, ***MS***
	“Liǎngyáng”	***YJ***
	“Wúhuà”	***WC***, HZ

Table 2.3: The Yuè dialect sites grouped according to the new classification

The main difference between this classification and Yue-Hashimoto's previous classification is that this classification does not subdivide the Pearl River Delta area in further detail. Yue-Hashimoto's granular classification is not supported by the measurements of lexical distance we observed. The distance between the various dialects of the Pearl River Delta area is dwarfed by the difference between those dialects and the dialects of the other groups of our classification. Internal classification of our Guǎngfǔ group might be possible, but likely not in the context of all the current data. There is still too much noise to divide any further for the current data set. Perhaps after more fieldwork has been conducted in the less documented regions, a more thorough analysis will be possible.

For now, this classification should be sufficient as a preliminary internal grouping. It accomplished to problem we set out to solve: to find the most significant internal groups in the Yuè dialect family. Further subdivision is not necessary for our purposes.

Chapter 3
HISTORICAL BACKGROUND

The immigration of Chinese speakers to modern Guǎngdōng and Guǎngxī started over two millenia ago. The first wave came with the Qín 秦 occupation of the south as military agriculturalists were forced to settle to defend the new territories. A second wave of migration occurred at the beginning of the Eastern Hàn 漢 dynasty, as refugees fled natural disasters and political turmoil in the north, firmly setting the Chinese identity of the region. A final large wave of Chinese immigration to this area occurred during the 12th and 13 centuries at the end of the Sòng period as the culmination of northerners gradually fleeing south to escape Mongol incursions.

3.1 The Qín Dynasty's Southern Campaign

The first major wave of Chinese migration to the region south of the Yángzǐ was during the Qín 秦 dynasty's southern campaign in 214BC. The Huáinánzǐ 淮南子 (HNZ) records some 500,000 people, split evenly into 5 forces settling in Yúgàn 餘干, Tánchéng 鐔成, Jiǔyí 九疑, Nányě 南野, and Pānyù 番禺 (HNZ 18:15a).

These people would likely have been dispatched from the Nányáng 南陽 basin in modern Hénán 河南 and followed the Yángzǐ 揚子 to the Xiāng 湘 and Gàn 贛 river basins. Those who followed the Xiāng river kept going south along the Guì River (Guì

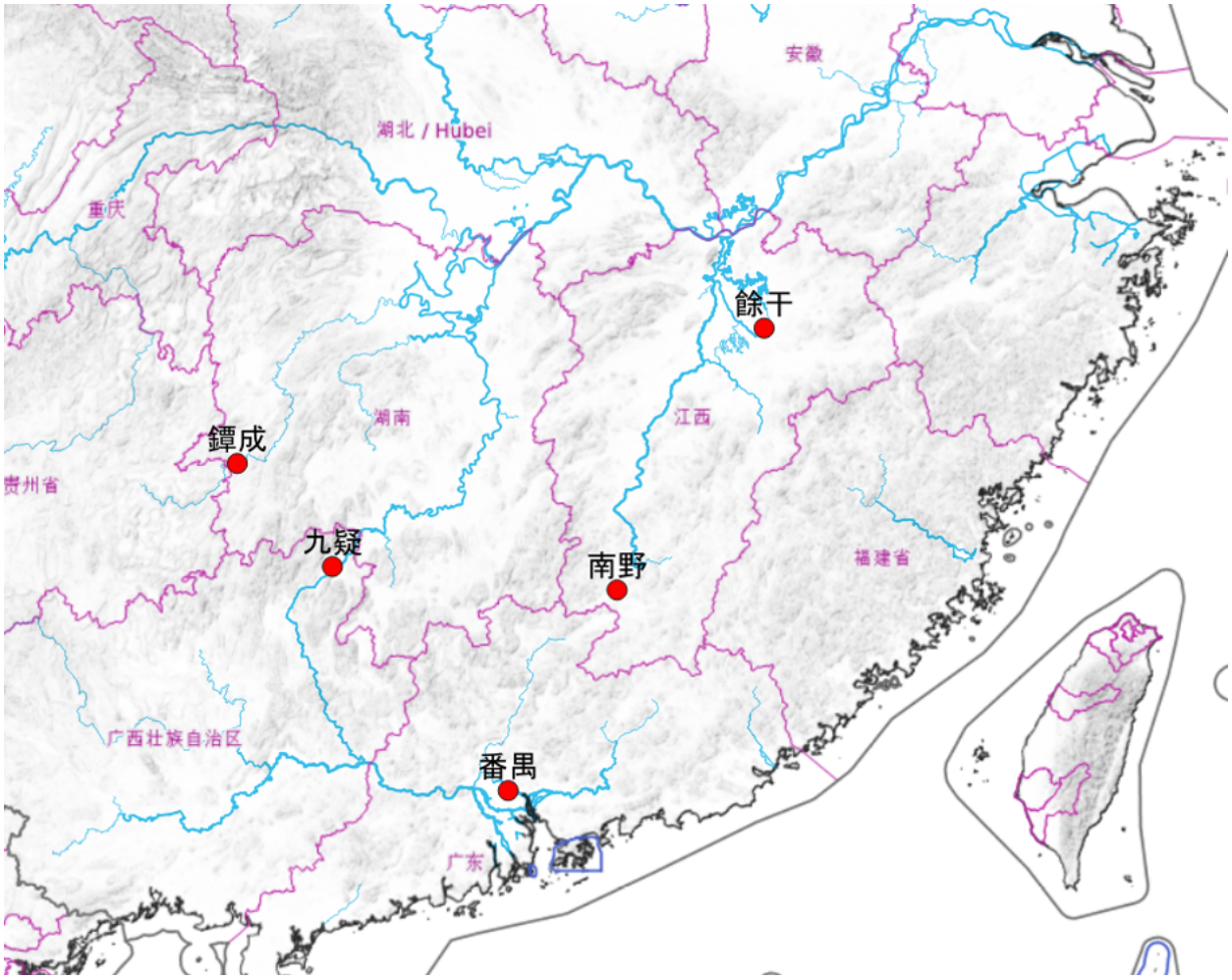


Figure 3.1: Qín Dynasty Southern Campaign Military Encampments

jiāng 桂江) to the West River and finally the Pearl River. They ended up near modern Guǎngzhōu 廣州. Each of the encampments are placed upon the major terminals of the river basin network. These people were not professional soldiers, but rather military agriculturalists intended to permanently settle the area. This is made explicitly clear in Shǐij 史記 (SJ). Sīmǎ Qiān 司馬遷 states that the Qín new commanderies in Guǎngdōng were made to accommodate the relocation of "deserters, freeloaders, and merchants" (SJ 6: 253, 113: 2967) The HNZ elaborates on this situation further. This relocation was intended to reinforce garrisons after more than 100,000 casualties while holding the Lǐngńan 嶺南 region (HNZ 18:15b). Diana Lary points out, many of these early Chinese settlers would have assimilated with the local population (2012, 39). Gě et al. estimate that less than 100,000 settlers were left in Guǎngdōng at the end of the campaign on the basis of a petition to Qín Shǐhuáng 秦始皇 that 30,000 women be sent to the region to marry the military agriculturalists (Gě et al. 1997, 2:73, SJ 118:3086). Only half that number was granted. The motivation for such a petition was likely that many of the settlers were "going native" and marrying into location populations. Basic vocabulary in Yuè dialects, especially in the case of body parts, animals, and fruit, suggests that creolization had begun to occur in some places. Many of these words are ultimately of Kra-dai origin.

When the Qín dynasty collapsed, Zhào Tuó 趙佗, previous commandant of Lóngchuān 龍川 county in the Nánhǎi 南海 commandery rose up and took control of the Qín's previous holdings in modern day Guǎngdōng, Guǎngxī, and Vietnam. It was not until

111BC that Hàn Wǔdì 漢武帝 managed to recapture the territory. Afterwards, up through the end of the Western Hàn 漢, emperors granted large numbers of fiefs in the Chángshā 長沙 commandery to better secure the channels into Guǎngdōng and the far southern holdings. Despite this policy, the population of the Southern border regions only grew modestly. Even if we assume that Qín Shǐhuáng was able to relocate enough people to replenish all casualties of his original force of 500,000 settlers, the population of the region only grew by an average of 4‰ annually to a total of 1,195,566. Gě (1 2002, p.344) estimates that the average national growth rate of this time was around 7‰ (2002, 1:344). However, Hans Bielenstein (1947, p.143) suggests that 4‰ would have been a natural growth rate for less agriculturally advanced settlements. It seems that from the original Qín campaign to the end of Western Hàn migration to the south was marginal and was likely cancelled out by gradual assimilation into the indigenous population.

3.2 Wáng Máng and the Second Wave

The Eastern Hàn saw a dramatic increase in migration to the southeastern territories. Bielenstein observed that the populations of modern Húnán, Jiāngxī 江西, and Guǎngdōng quadrupled during the time between the 2AD census recorded in Hànshū (HS) 漢書 and the census of 140AD recorded in Hòuhànshū (HHS) 後漢書, leading him to remark "such an enormous increase cannot be due to nativity alone" (Bielenstein, 1947, p.141). Gě et al. agree that at least modern Húnán and Jiāngxī saw appreciable population growth due to migration during this time but are more skeptical that southward migration into

this region extended as far as the immediate Lǐngnán region (2 Gě et al., 1997, pp.262–65) (1997, 2:262–65). Gě (1 2002, p.421) also casts doubt upon whether the 140AD census data is completely accurate. Many of the princes holding fiefs in Chángshā at the time would likely have inflated their counts to garner more support from the central government. However, he estimates this inflation might ultimately be insignificant since a great portion of the population of those regions may not have been registered for tax collection (2002, 1:421). The apparent increase in population to an extent is probably partially due to greater registration, but in the case of the south, cannot account for the dramatic change alone. The total amount of migrants involved in this growth cannot be understated. Looking at the Yùzhāng 豫章, Línglíng 零陵, Guìyáng 桂陽, Nánhǎi, Cāngwú 蒼梧, Yùlín 鬱林, and Hépǔ 合浦 commanderies, we see a range of growth rates ranging from 7.1‰ to 14.3‰, all above the nation average. This rate of growth is even more explosive when compared to neighboring regions that did not connect to the Gàn or Xiāng river basins. These range from 1.4‰ to 3.4‰. If we assume that the growth of the Early South Central Chinese commanderies would have been near the nation average of 7‰, then at least 1.5 million people would have had to migrate to the region between 2AD and 130AD. If we assume the old Western Hàn growth rate of 4‰, then the total goes up to more than 2.3 million. This large wave of migration must have greatly influenced the language spoken in these regions and its effects likely linger on in residually in Common Yue.

Bielenstein (1947, p.141) estimates that the original catalyst of this migration wave

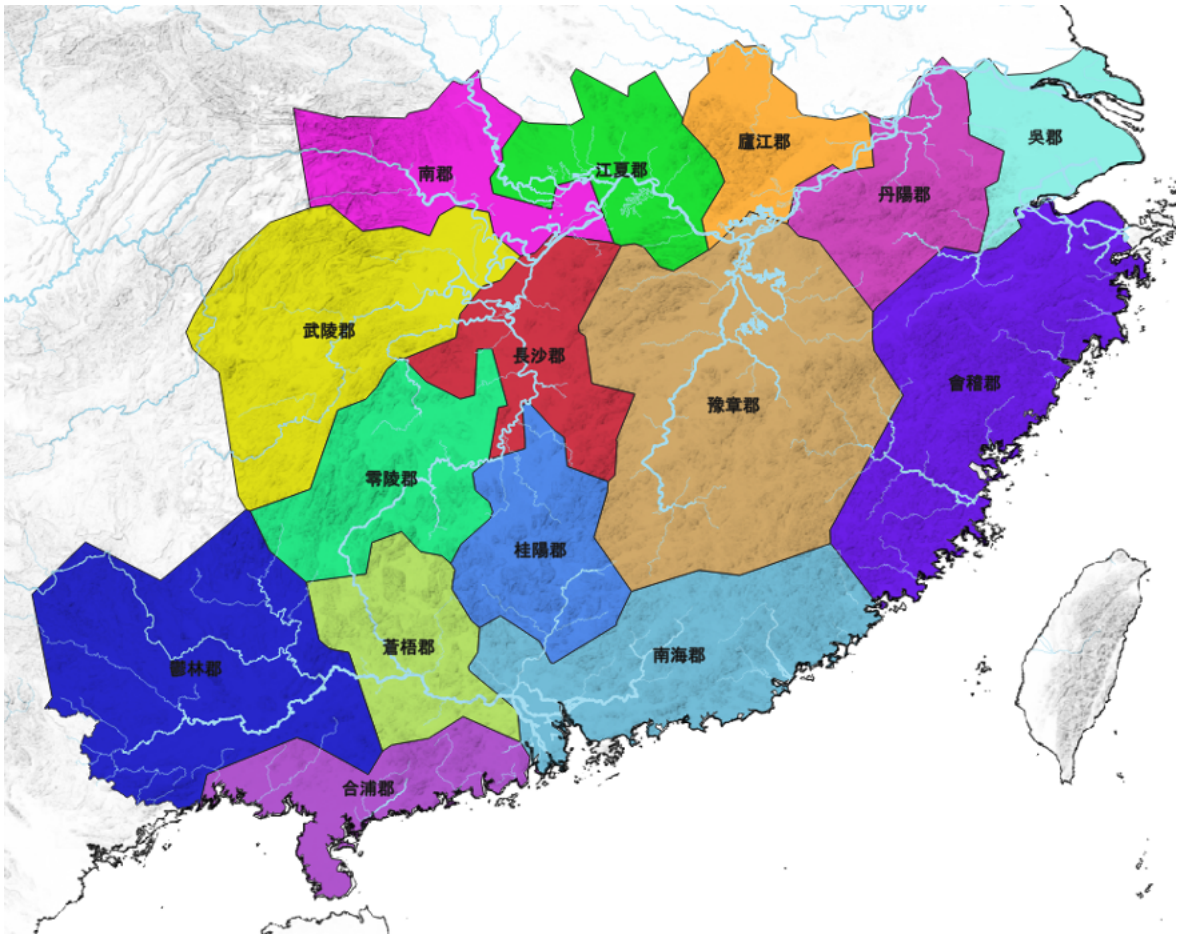


Figure 3.2: Southern Commanderies of the Eastern Hàn

Commandery	HS pop.	HHS pop.	Rate	7‰ Est.	4‰ Est.
Yùzhāng 豫章郡	351965	1668906	11.3‰	747255	1058313
Línglíng 零陵郡	139378	1001578	14.4‰	636605	759783
Guìyáng 桂陽郡	156488	501403	8.5‰	91626	229926
Nánhǎi 南海郡	94253	250282	7.1‰	3472	86771
Cāngwú 蒼梧郡	146160	466975	8.5‰	84242	213415
Yùlín 鬱林郡	71162	(no data)			
Hépǔ 合浦郡	78980	86617	0.06‰		
Wǔlíng 武陵郡	157180	250913	3.4‰		
Dānyáng 丹陽郡	405170	630545	3.2‰		
Wújùn 吳郡	516295	700782	2.2‰		
Jiāngxià 江夏郡	219218	265464	1.4‰		
Lújiāng 廬江郡	457333	424683	0‰		

Table 3.1: Population of Southern Commanderies According to HHS with estimates of total migrants to each given 7‰ and 4‰ estimates of base population growth

was dykes on the Yellow River breaking in 11AD, causing mass flooding over the area. This combined with the chaos of Wáng Máng's 王莽 rebellion, likely drove many people southward. Curiously, not much explicit documentation of this migration remains in the official histories. This is likely because the majority of the migrants were unlettered common people and agriculturalists. However, there are still a few clues outside of the census data that corroborate this migration.

Wáng Máng appears to have continued the Western Hàn practice of relocating nobles to Chángshā. He renamed the Chángshā commandery to Tiǎnmán 填蠻 "Filling in the Mán (Barbarians)." This alone tells us much about his policy. Like his predecessors, he was concerned about losing connections with the far south territories and wanted to encourage Chinese citizens to fill in towns along the Xiāng riverways. There are little administrative

records of his short-lived Xīn 新 dynasty, but it is conceivable that at least some of the relocation to modern Húnán at this time was under state coercion.

Although not many new counties were established in the commanderies discussed above between 2AD and 140AD, the few that were stand out in terms of location and time. In Guǎngdōng, Zhāngpíng 鄣平 county was established in 71AD at the far west side of the Cāngwú commandery, now along the western border with Guǎngxī, and Zēngchéng 增城 county was carved out the northeast portion of Pānyù 番禺 county in 136AD to account for the growing population. Immediately northwest of modern Guǎngdōng, in modern Húnán, Hànníng 漢寧 county was established in 136AD in the middle of the Guìyáng commandery (HHS 96:3483). These new administrative regions demonstrate an influx of population around the end of the first century AD into the Guǎngdōng region. These migrants would have gradually moved there over the decades since their original foray south of the Yángzǐ at the beginning of the century.

Finally, many original counties of the Chángshā commandery were reapportioned to the neighboring southern Línglíng Commandery during the first half of the Eastern Hàn dynasty. This was likely in response to the growing population of the southern-most regions of Chángshā. They had grown to a point that they needed to be accounted for by a new administrative zone.

3.3 Migration in the Sòng Period

Gě et al. (3 1997, p.174) lament the lack of historical sources and scholarship on immigration to Guǎngdōng in the Sòng period. Censuses in official histories record dramatic rises and falls of individual counties, making it difficult to contextualize this data regarding immigration. Although largely recorded in the Qīng dynasty, the family lineages of Guǎngfǔ group speaking people help paint a slightly clearer picture. Many of these lineages record that families originally came from Nánxióng 南雄, located in the northern most part of modern Guǎngdōng northwest of Sháoguān 韶關, during the Sòng period. Gě et al. (3 1997, pp.179–181) observes that these accounts agree with historical accounts of loyalists and military deserters making their way south to Nánxióng at the end of the 13th century after the fall of Zhèjiāng 浙江 to the Mongols. It is likely that the northern *koiné* that informs those most recent literary layers of Common Gan, Common Neo-Hakka, and Common Yue ultimately comes from the northern refugees of this migration wave.

However, the effect of this migration is probably not just limited to the literary layers of Common Yue. The distribution of Yuè dialects, as laid out in Chapter 2, suggests a clear route of migration to the area. We do not find Yuè dialect distributed along the Guì and Liǔ 柳江 rivers or elsewhere in the north of Guǎngxī province. Rather, we find them on the North river around Nánxióng and Sháoguān, stretching down into the Pearl River Delta, then westward along the West, Xún, and Yōng rivers in the sout of Guǎngxī. This suggests that the initial wave of Common Yue speakers originated north of Nánxióng

and spread south and westward along the major waterways. This suggests they shared a homeland with the speakers of Coblin's Early Southern Highlands Chinese and probably also share older layers of lexicon that had accumulated in the vernacular of that region before the more recent introduction of Sòng dynasty koinés.

Therefore, we propose that Common Yue represents a multi-stratal phonological common system that spread outward from northeastern Guǎngdōng in waves likely beginning in the 12th century. However, it also contains some substrates of vocabulary from before this expansion from the original population that occupied Guǎngdōng and Guǎngxī before this migration. It is possible that some of these linguistic substrates are shared by dialects in northern Guǎngxī since the migration route of this earlier stage of settlement was likely out of Húnán and through northern Guǎngxī into the Lǐngnán area. Comparisons with Coblin's Early Southern Highlands Chinese and the phonological common systems that underly the northern Guǎngdōng and Guǎngxī patois and northern Pínghuà dialects will no doubt shed more light on these connections and migration waves.

Chapter 4

PHONOLOGICAL OVERVIEW

Fieldwork on Chinese dialects is traditionally recorded in narrow phonetic transcriptions. Phonological analysis is typically limited to the complementary distribution of consonant onsets and tone, while vowels are usually left entirely untouched. Sometimes, investigators will try to restrict vowel phonemes to a limited set, but always separate out what they perceived to be medial elements first according to native philological traditions. They never treat sequences of multiple vowels as complex vowel nuclei. While this presentation has value in portraying the recorded dialects candidly in some cases, it is not an optimal starting place for diachronic comparison. Narrow transcriptions will often obscure regular correspondence across dialects and can render the notation used to present common innovations unwieldy. It also casts allophones of a single phoneme as separate segments, thereby complicating the stage of the comparative method in which segments in complementary distribution are merged. In this study, we will reduce the original narrow transcription used in the fieldwork we draw upon to a broad phonological transcription to more easily facilitate comparison and discourse. This analysis is intended to be functional, not necessarily the most optimal synchronic analysis.

4.1 Organization of Correspondences and Syllable Structure

Historical linguists on the whole commonly organize correspondences in terms of phonemic segments. In Chinese historical linguistics, correspondences are more commonly organized into syllable "initials," consisting of an optional simple consonant onset; "finals," consisting of an option vocalic on-glide, an obligatory syllable vocalic nucleus, an optional vocalic off-glide, and an optional consonant coda; and tones. This sort of organization can be advantageous because individual phonemic segments are overwhelmingly morpho-phonemically static across varieties of Chinese. Analyzing these segments in the same terms one might use for Indo-European languages, for example, looking at "medial" and "initial" environments of consonant segments, is not useful in the case of Chinese.

However, the organization of correspondences into traditional Chinese initials and finals can suffer from the clumsiness of the sheer number of correspondence sets that must be established to cover all possible finals. Yuè dialects are rich in consonant codas and vocalic nuclei, making the total number of distinctive Common Yue finals quite high. To afford our discussion more economy, we will take a compromise approach and organize our correspondences into onset, nucleus, coda, and tone divisions of the syllable. We shall call these divisions "components" to avoid confusion of terminology with phonemic "segments." The onset and coda components of the syllable constitute the "margins" of the syllable and are optional. The nucleus of the syllable is obligatory. Tones are suprasegmental, and while typically obligatory in Chinese morphemes, they can float

across multiple syllables in some varieties of Chinese.

4.1.1 Tone

Zhān, Cheung, et al. (1987a,b; 1994; 1998), Xiè (2007), and Xiǎn (2016) and Lǐ (2014) write out tone contours phonetically using Chao tone numbers. We convert these to single digits using the phonemic tone numbers of Common Dialect Chinese (CDC) (Norman, 2006). Additionally, the phonemic upper *Yīnrù* yone (*shàng yīnrù* 上陰入) is marked as 7 while the lower *yīnrù* tone (*xià yīnrù* 下陰入) is marked as 9. The same is done, mutatis mutandis, for the upper *Yángrù* yone (*shàng yángrù* 上陽入) tone, 8, and the lower *yángrù* tone (*xià yángrù* 下陽入), 10.

	Píng 平	Shàng 上	Qù 去	Rù 入
Yīn 陰 Register	1	3	5	7
Yáng 陽 Register	2	4	6	8

Table 4.1: CDC Tone Notation

Morphologically “changed tones” are written with the formula T-XY, where T is the underlying phonemic CDC tone number and XY are phonetic Chao tone numbers. In cases where the underlying phonemic tone is not clear synchronically due to a lack of morphemic variation, we simply omit the T. Tones changed as a result of regular tone sandhi are left unmarked except in cases where the underlying tone is clear from the surface form. In cases where the underlying tone is unclear, we apply the same convention as the morphemic tone changes.

The 53 and 55 variants of tone 1 in GZ are commonly analyzed as allophones (Yue-Hashimoto, 1972; Bauer and Benedict, 2011). It is reported that in the speech of older GZ speakers, 55 is the form tone 1 takes before other tones that begin in a high pitch, while its citation tone is 53 elsewhere. In the speech of younger people, this allophonic distribution has been disrupted. We have not observed such a distribution in the surveys on the Guǎngfǔ groups in this study, so we always explicitly write out either 1-55 or 1-53 in cases of dialects where both variants exist.

4.1.2 Onset versus Initial

The "initial" of the traditional analysis of Chinese syllable structure is similar to the "onset" of the universal model of the syllable, with one important difference. Syllable onsets may be complex and may be licensed vocalic features, like [front/palatal], [round/labial] and [back/velarized] (Kehrein and Golston, 2004; Golston and Kehrein, 2015). Therefore, "medial" on-glides of the finals of traditional Chinese analyses may in some cases be "moved" in the analysis to be constituents of the onset. There is a great amount of debate over whether "medials" of Chinese dialects are phonologically part of onsets or nuclei. On the extreme end of the spectrum, Duanmu (2007) argues that "medials" maybe interpreted as part of the onset in all Chinese dialects. More moderate proposals argue for different analyses on a dialect-by-dialect basis (Lin, 1989; Bāo, 1996).

In our study of Yue dialects, we follow Maddieson (1984) as a general model. He proposes that there are three possible phonological interpretations of diphthongs (Mad-

dieson, 1984, p.161):

1. As a phonemic unit
2. As a vowel and consonant in sequence (in either order)
3. As two vowels in sequence (i.e., part of different syllables)

The third interpretation is never applicable to the phonetic diphthongs and triphthongs of Yue dialects. These phones always occur within single syllables. For the other two interpretations, Maddieson suggests that the first interpretation can be adopted when "phonological unity is indicated by the fact that [diphthongs] have the same distributional patterns as simple vowels with respect to syllables and tones," while the second analysis can be adopted when part of a diphthong may be reinterpreted as a consonant that fits the pattern established by other consonants in syllable margins (Maddieson, 1984, p.161). To illustrate the second case, he raises the example of Standard Thai. Phonetic diphthong final elements [i] and [u] in Thai do not occur before consonant endings. They fit the pattern of other syllable-final sonorant consonants, so they are analyzed as sequences of vowels followed by approximant consonants /-j/ and /-w/.

In the case of GZ Cantonese, many scholars have proposed that the onglide "medial" [u] is in fact part of the onset because it only occurs with zero or velar "initials" (Yue-Hashimoto, 1972; Light, 1977; Yip, 1997). We adopt this interpretation for the majority of Yue dialects in the study and rewrite cases of [u] following zero and velar initials and preceding other vowels as labiovelars onsets /w-, k^{w}-, k^{wh}-, $ŋ^{w}$-/, etc. This new set of consonants exhibits phonological restrictions similar to the phonological restrictions of

bilabial onsets and does not occur in syllables that also end in labial finals, i.e. no there is no /pap/ or /kʷap/. Exceptions to this general rule are the nuclei /uə/ in YS, /ua/ in EP, KP, and MS that do not exhibit similar patterns of restriction. They occur with other initials besides those that are zero and velar. Therefore, we treat them as phonemically unitary diphthongs.

The case of [i] on-glides in Yue dialects is less straight-forward. In all cases where such on-glides are found after zero initials in the traditional analysis, they are reanalyzed simply as the onset /j-/. Zhān, Cheung, et al. (1987a,b; 1994; 1998) often add [(j)] before such "medials" in their transcription. They do not explain why they add the parentheses, but presumably it is because they believe the [j] segment is not phonologically independent. In their transcriptions, zero initials do not occur before high front vowels, only [j]. We rewrite these cases as simple /j/. In cases where [i] on-glides follow initials that are not zero or [j], they usually only occur before a limited set of other vowels, typically only [a] and [ɔ]. We analyze them as part of phonemically unitary diphthongs serving as syllable nuclei. True phonemically unitary diphthongs are typologically limited to just a small number per phonemic inventory cross-linguistically (Maddieson, 1984). Cases where "medials" are limited so that they consistently occur before just one or two other vowels are probably best analyzed as phonemic units.

In some cases, [i] is otherwise analyzed as part of an allophonic expression of other simple nuclei, e.g. [-ia-] is as an allophone of /-ɛ-/ in some dialects. Occasionally, on-glide [i] follows coronal nasal onsets in Zhān, Cheung, et al. (1987a,b; 1994; 1998) where

such an [i] is not otherwise part of a diphthong nucleus. These cases are a minority. We reanalyze the [ni] sequences as palatal nasal onsets and cover their exact distribution in the discussion of the various dialects below.

4.1.3 Nucleus and Coda versus Final

The ending subdivision of the traditional "final" and the coda syllable component are again similar. In the traditional Chinese analysis, off-glides are considered part of the nucleus, but in actual practice, they overwhelmingly compete with other consonants to occupy the coda slot. To our knowledge, only some Northern Min dialects permit off-glides in nuclei before consonant codas, like in [-eiŋ], [-øyŋ], etc. In almost all other varieties of Chinese, off-glides are only permitted when other consonant codas are absent. The same vowels that occur before other sonorant consonant codas also occur before final [i] and [u]. In other words, they exhibit the same behavior Maddieson observed in Thai. As such, we adopt his treatment. To make this interpretation clear, we shall rewrite the off-glides [-i] and [-u] as /-j/ and /-w/ respectively to indicate they are phonemic codas.

As stated above, the remaining cases of the on-glide [i] in "finals" that cannot otherwise be analyzed as part of the onset are treated as part of unitary diphthongs in the nucleus position. There are no other cases of other "medial" on-glides and complex nuclei as a whole are limited in distribution across Yuè dialects.

4.2 Other Considerations

We have elected to rewrite all nonstandard IPA symbols found in the dialect descriptions in standard IPA. Documenters of Chinese dialects, especially those with more traditional training backgrounds from mainland China, tend to use now obsolete phonetic symbols in their transcriptions. These idiosyncrasies may easily be converted to standard IPA without any loss of information and their elimination would likely help this work appeal to a wider audience.

In some cases, there is only limited or otherwise suspicious evidence of minimal pairs to establish phonemic contrast between syllabic components. We prefer to analyze these instances as nondistinctive, but use parentheses to preserve the original transcription so as not to erase distinctions made by the authors. For example, in the phonological analysis of FK dialect given by Zhān, Cheung, et al. (1998), they explicitly state that the segments [ni] and [n̩] do not contrast and opt to write them all as [ni]. However, in their listing of the data, they use both transcriptions and even record minimal pairs. We follow Zhān, Cheung, et al. (1998)'s prose explanation and suspect that these cases are due to errors in editing and do not represent true contrast. However, to preserve the original contrast for readers, we represent [ni] as /ɲ/ and [n̩] as /ɲ(ʲ)/. We discuss all other cases of transcription preserving amendments to our alterations below.

Transcriptions of character readings and transcriptions of characters used to represent actual words in the lexical studies often conflict with each other. Such conflicts are mostly

phonemic in nature and differences can be represented using our new broad transcription. For example, in NH the character reading of 玩 is listed as [fun^{22}], but in the word for "to play" in the vocabulary comparison, also written 玩, it is transcribed as [p^{h}aŋ^{35}]. This sort of mismatch is widespread in the Zhān, Cheung, et al. (1987a,b; 1994; 1998) and Zhān et al. (2000) surveys, but is relatively less common in the other surveys. In Xiè (2007), character readings almost never deviate from the readings given in the transcription of vocabulary and very few characters have multiple readings. Both results are suspicious and raise serious questions about their methods of elicitation. Ideally, characters should have multiple readings to cover all the ways they are pronounced in the words they write. In Yuè dialects, we expect at least two readings in a nontrivial number of characters to reflect a clear separation of "literary" and "vulgar" readings.

Sub-phonemic conflicts in transcription also exist. For example, in TS the character 飢 is listed with the reading [kei^{33}], but in the transcription of the word for "hungry" it is written 肚飢 [u^{55} ki^{33}]. In a more extreme example, the finals [-øŋ] and [-øk] in the character readings of the TS dialect never occur in the vocabulary comparison. They are instead written [-oŋ] and [-ok]. We try to preserve these differences in notation while also explicitly marking them as nondistinctive using parentheses as mentioned above. Generally, we favor the transcriptions in the vocabulary comparisons since they inform the basis of our correspondences. All quoted pronunciations following orthography are informed by the lexical studies unless it is explicitly stated that they are character readings.

In the table below, we give the original transcriptions on the left and our new tran-

scriptions on the right separated by a slash.

4.3 The Fēnglián Group

There appears to be regressive tone sandhi in the YS dialect, but sadly Zhān, Cheung, et al. (1994) dedicate no discussion to it. We observed tones 1, 2, 3, 5, and 7 surfacing as [^{33}] in nonfinal position in lexical compounds. However, tone 1 also seems to surface as [^{42}] and sometimes [^{31}] before another tone 1 and tone 5 as [^{24}] before another tone 5. The tone 1 changes pose no surface distinction issues, but the rest pose phonemic ambiguity, so we will write out these values after the citation tone.

p/p	p’/p^{h}	m/m	f/f					
t/t	t’/t^{h}	n/n		l/l				
ts/ts	ts’/ts^{h}		s/s	j,(j)i/j				
k/k	k’/k^{h}	ŋ/ŋ	h/h					
ku/k^{w}	k’u/k^{wh}	ŋu/ŋ^{w}		w/w				
a/a		œ/œ	ɔ/ɔ	e/e	o/o	i/i	u/u	y/y
ai/aj	ɐi/ɐj	ɔy/œj			oi/oj		ui/uj	
au/aw	ɐu/ɐw			eu/ew		iu/iw		
an/an	ɐn/ɐn			en/en	on/on	in/in	un/un	yn/yn
	ɐŋ/ɐŋ	œŋ/œŋ	ɔŋ/ɔŋ	eŋ/eŋ	oŋ/oŋ	iŋ/iŋ		
am/am	ɐm/ɐm							
at/at	ɐt/ɐt			et/et	ot/ot	it/it	ut/ut	yt/yt
	ɐk/ɐk	œk/œk		ek/ek	ok/ok	ik/ik		
ap/ap	ɐp/ɐp							
m̩/m̩								
52,42,31/1	55/3	35/5	55/7	33/9				
342/2	24/4	22/6	22/8					

Table 4.2: YS Transcription Comparison

The redundant palatal nasal notation of FK was already covered. In addition, in the listing of finals Zhān, Cheung, et al. (1994) give FK /ɔ/ as [o], but otherwise write it as

p/p	p’/p^{h}	m/m	f/f					
t/t	t’/t^{h}	n/n		l/l				
ts/ts	ts’/ts^{h}		s/s					
tʂ/tʃ	tʂ’/tʃ^{h}	ɳ/ɲ	ʂ/ʃ	j,(j)i/j				
k/k	k’/k^{h}	ŋ/ŋ	h/h					
ku/k^{w}	k’u/k^{wh}			w,(w)u/w				
a/a		ɛ/ɛ	œ/œ	ɔ/ɔ	i/i	u/u	y/y	
ai/aj	ɐi/ɐj		œy/œj	ɔj/ɔj		ui/uj		
au/aw	ɐu/ɐw	ɛu/ɛw		o/ɔw	iu/iw			
an/an	ɐn/ɐn	ɛn/ɛn		on/ɔn	in/in	un/un	yn/yn	
aŋ/aŋ		ɛŋ/ɛŋ	œŋ/œŋ	ɔŋ/ɔŋ	iŋ/iŋ		yŋ/yŋ	iaŋ/iaŋ
at/at	ɐt/ɐt	ɛt/ɛt		ot/ɔt	it/it	ut/ut	yt/yt	
ak/ak		ɛk/ɛk	œk/œk	ɔk/ɔk	ik/ik		yk/yk	iak/iak
m̩/m̩								
53/1	55/3	33/5	55/7	33/9				
21/2	35/4	22/6	22/8					

Table 4.3: LX Transcription Comparison

[ɔ] in actual transcriptions. We adopt the latter. When [i] and [u] occur with a zero initial before another vowel, we rewrite them as /j-/ and /w-/. Finally, FK also exhibits regressive tone sandhi. In compounds, all syllables before the final syllable can change depending on their underlying tone. Zhān, Cheung, et al. (1998) does not give explicit rules for this sandhi. According to our observations, tones 1, 3, 5, and 9 surface as [^{33}] and tones 2, 4, and 7 surface as [^{22}] in compounds in nonfinal position. We write these out explicitly since they can cause ambiguity.

p/p	p’/p^{h}	m/m	f/f				
t/t	t’/t^{h}	n/n	θ/θ	l/l			
ts/ts	ts’/ts^{h}			j,(j)i/j			
k/k	k’/k^{h}	ŋ/ŋ	h/h				
ku/k^{w}	k’u/k^{wh}			w/w			
a/a		ɛ/ɛ	œ/œ	ɔ/ɔ	i/i	u/u	y/y
ai/aj	ɐi/ɐj		œy/œj	ɔi/ɔj		ui/uj	
au/aw	ɐu/ɐw	ɛu/ɛw		ɔu/ɔw	iu/iw		
an/an	ɐn/ɐn	ɛn/ɛn		ɔn/ɔn	i/in	un/un	yn/yn
aŋ/aŋ	ɐŋ/ɐŋ	ɛŋ/ɛŋ	œŋ/œŋ	ɔŋ/ɔŋ		oŋ/uŋ	
am/am	ɐm/ɐm				im/im		
	ɐt/ɐt	ɛt/ɛt		ɔt/ɔt	it/it	ut/ut	yt/yt
ak/ak		ɛk/ɛk	œk/œk	ɔk/ɔk	ik/ik	ok/uk	
	ɐp/ɐp				ip/ip		
m̩/m̩	ŋ̍/ŋ̍						
42/1	53/3	33/5	55/7	33/9			
231/2	13/4	325/6	325/8				

Table 4.4: HJ Transcription Comparison

p/p	p’/p^{h}	m/m	f/f				
t/t	t’/t^{h}	n/n		l/l			
ts/ts	ts’/ts^{h}	ni,ɲ/ɲ(^{j})	s/s	i/j			
k/k	k’/k^{h}	ŋ/ŋ	h/h				
ku/k^{w}	k’u/k^{wh}			u/w			
a/a		e/e	œ/œ	ɔ/ɔ	i/i	u/u	y/y
ai/aj	ɐi/ɐj	ei/ej	œy/œj			ui/uj	
au/aw	ɐu/ɐw	eu/ew		ou/ɔw	iu/iw		
am/am	ɐm/ɐm	em/em			im/im		
an/an	ɐn/ɐn	en/en		ɔn/ɔn	in/in	un/un	yn/yn
aŋ/aŋ	ɐŋ/ɐŋ	eŋ/eŋ	œŋ/œŋ	oŋ/ɔŋ	iŋ/iŋ	uŋ/uŋ	iuŋ/yŋ
ap/ap	ɐp/ɐp	ep/ep			ip/ip		
at/at	ɐt/ɐt	et/et		ɔt/ɔt	it/it	ut/ut	yt/yt
ak/ak	ɐk/ɐk	ek/ek	œk/œk	ok/ɔk			
					iə	uə	
ŋ̍/ŋ̍							
55/1	52/3	433/5	55/7	43/9			
24/2	242/4	21/6	22/8				

Table 4.5: FK Transcription Comparison

4.4 The Guǎngfǔ Group

The DG dialect only has the glottal stop coda once after /a/. It is in complementary distribution due to tone class distinctions, so we omit it to keep the coda places of articulation consistent.

p/p	p'/p^{h}	m/m	f/f				
t/t	t'/t^{h}	n/n		l/l			
ts/ts	ts'/ts^{h}		s/s	j,(j)i/j			
k/k	k'/k^{h}	ŋ/ŋ	h/h				
ku/k^{w}	k'u/k^{wh}			w/w			
a/a		ɛ/ɛ	œ/œ	ɔ/ɔ	i/i	u/u	y/y
ai/aj	ɐi/ɐj		œy/œj	ɔi/ɔj	ei/ij	ui/uj	
au/aw	ɐu/ɐw				iu/iw	ou/uw	
an/an	ɐn/ɐn		œn/œn	ɔn/ɔn	in/in	un/un	yn/yn
aŋ/aŋ	ɐŋ/ɐŋ	ɛŋ/ɛŋ	œŋ/œŋ	ɔŋ/ɔŋ	eŋ/iŋ	oŋ/uŋ	
am/am	ɐm/ɐm				im/im		
at/at	ɐt/ɐt		œt/œt	ɔt/ɔt	it/it	ut/ut	yt/yt
ak/ak	ɐk/ɐk	ɛk/ɛk	œk/œk	ɔk/ɔk	ek/ik	ok/uk	
ap/ap	ɐp/ɐp				ip/ip		
m̩/m̩	ŋ̍/ŋ̍						
53,55/1	35/3	33/5	55/7	33/9			
21/2	13/4	22/6	22/8				

Table 4.6: GZ Transcription Comparison

p/p	p’/p^{h}	m/m	f/f				
t/t	t’/t^{h}	n/n		l/l			
ʈ							
ts/ts	ts’/ts^{h}		s/s	j,(j)i/j			
k/k	k’/k^{h}	ŋ/ŋ	h/h				
ku/k^{w}	k’u/k^{wh}			w/w			
a/a		ɛ/ɛ	œ/œ	ɔ/ɔ	i/i	u/u	y/y
	ɐi/ɐj		œy/œj	ɔi/ɔj		ui/uj	
au/aw	ɐu/ɐw	ɛu/ɛw			iu/iw	ou/uw	
	ɐn/ɐn		œn/œn	ɔn/ɔn	iŋ/in	un/un	yn/yn
aŋ/aŋ		ɛŋ/ɛŋ	œŋ/œŋ	ɔŋ/ɔŋ	eŋ/iŋ	oŋ/uŋ	
am/am	ɐm/ɐm	ɛm/ɛm			im/im	om/um	
	ɐt/ɐt	ɛt/ɛt	œt/œt	ɔt/ɔt	ik/it	ut/ut	yt/yt
ak/ak		ɛk/ɛk	œk/œk	ɔk/ɔk	ek/ik	ok/uk	
	ɐp/ɐp	ɛp/ɛp			ip/ip	op/up	
m̩/m̩	ŋ̍/ŋ̍						
55/1	35/3	33/5	55/7	33/9			
44/2	13/4	22/6	22/8				

Table 4.7: NH Transcription Comparison

p/p	p’/p^{h}	m/m	f/f	v/v			
t/t	t’/t^{h}	n/n					
ts/ts	ts’/ts^{h}		s/s	z/z			
k/k	k’/k^{h}	ŋ/ŋ	h/h				
ku/k^{w}	k’u/k^{wh}						
a,aʔ/a		ɛ/ɛ	ø/ø	ɔ/ɔ	i	u	y
ai/aj	ɐi/ɐj			ɔi/ɔj		ui/uj	
au/aw				ɔu/ɔw	iu/iw		
	ɐm/ɐm						
	ɐn/ɐn		øn/øn		in/in	un/un	
aŋ/aŋ		ɛŋ/ɛŋ	øŋ/øŋ	ɔŋ/ɔŋ	əŋ/iŋ	oŋ/uŋ	
	ɐp/ɐp				ip/ip		
	ɐt/ɐt		øt/øt		it/it	ut/ut	
ak/ak	ɐk/ɐk	ɛk/ɛk	øk/øk	ɔk/ɔk	ək/ik	ok/uk	
m̩/m̩							
213/1	35/3	32/5	44/7	224/9			
21/2	13/4		22/8				

Table 4.8: DG Transcription Comparison

p/p	p’/p^{h}	m/m	f/f				
t/t	t’/t^{h}	n/n	θ/θ	l/l			
ts/ts	ts’/ts^{h}		s/s	#i/j			
k/k	k’/k^{h}	ŋ/ŋ	h/h				
ku/k^{w}	k’u/k^{wh}			w/w			
a/a		ɔ/ɔ	e/e	ø/ø	i/i	u/u	y/y
ai/aj	ɐi/ɐj			øy/øj		ui/uj	
au/aw	ɐu/ɐw				iu/iw		
an/an	ɐn/ɐn	ɔn/ɔn			in/in	un/un	yn/yn
aŋ/aŋ	əŋ/ɐŋ	ɔŋ/ɔŋ		øŋ/øŋ	iəŋ/iŋ	oŋ/uŋ	
am/am	ɐm/ɐm				im/im		
at/at	ɐt/ɐt	ɔt/ɔt			it/it	ut/ut	yt/yt
ak/ak	ək/ɐk	ɔk/ɔk		øk/øk	iək/ik	ok/uk	
ap/ap	ɐp/ɐp				ip/ip		
ˈṅ/ṅ							
33/1	35/3		55/7	33/9			
22/2	242/4	21/6	22/8				

Table 4.9: FG Transcription Comparison

p/p	p’/p^{h}	m/m	f/f				
t/t	t’/t^{h}	n/n		l/l			
ts/ts	ts’/ts^{h}		s/s	j,(j)i/j			
k/k	k’/k^{h}	ŋ/ŋ	h/h				
ku/k^{w}	k’u/k^{wh}			w/w			
a/a		ɛ/ɛ	œ/œ	ɔ/ɔ	i/i	u/u	y/y
ai/aj	ɐi/ɐj		œy/œj	ɔi/ɔj	ei/ij	ui/uj	
au/aw	ɐu/ɐw				iu/iw	ou/uw	
(an/an)	ɐn/ɐn				in/in	un/un	yn/yn
aŋ/aŋ		ɐŋ/ɛŋ	œŋ/œŋ	ɔŋ/ɔŋ	eŋ/iŋ	oŋ/uŋ	
am/am	ɐm/ɐm				im/im		
(at/at)	ɐt/ɐt				it/it	ut/ut	yt/yt
ak/ak		ɐk/ɛk	œk/œk	ɔk/ɔk	ek/ik	ok/uk	
ap/ap	ɐp/ɐp				ip/ip		
m̩/m̩	ṅ/ṅ						
55/1	35/3	33/5	55/7	33/9			
21/2	13/4		22/8				

Table 4.10: YF Transcription Comparison

4.5 The Sìyì Group

The final [iat] in XH only occurs after the initial [z]. There is no [ziak] to constrast with it, so we reanalyze [iat] as /æk/. Where [i] occurs in medial position after [ᵑg] it is reinterpreted as part of the onset, /ŋʲ/.

TS [ei] and [i] are likely in complementary distribution, even though the character data lists minimal pairs for virtually all initials except [ɬ-]. There is also only one instance of [tsei] contrasting with an overwhelming representation of [tsi]. In all other published descriptions of subdialects of TS consulted, these finals do no contrast phonemically (Wáng and Qián, 1950; McCoy, 1966; Yue-Hashimoto, 2005). The distinction between the two does not correspond to a similar distinction in any other Yuè dialect, nor does it accord consistently with traditional QYS categories. Superficially, there appears to be some regularity in that members of the QYS *-je* 支 rime group that begin with bilabial initials and [l-] are consistently [i], while other similar rime groups, *-ij* 脂 and *-i* 之, with like initials are [-ei]. However, this is an artifact of how the data is laid out. The *-je* 支 rime group is listed first. Halfway through the group on page 93, the dominant transcription of the final switches to [ei]. Members of all three rime groups with velar initials are all [Kei] except for QYS-irregular 痣 [ki]. We suspect this is an issue in editing rather than the manifestation of some preserved archaic features. To preserve the original notation, all cases of [ei] are written as /ij/, despite a dubious contrast of it and /i/. The finals -øn (戀) and -øt (律率栗) only occur after the initial l- in Zhān, Cheung,

et al. (1987b) in character readings. They do not appear in any other descriptions of TS phonology and never occur in the transcriptions of Zhān, Cheung, et al. (1987a)'s lexical study. These are likely Guǎngfǔ-influenced character readings acquired by the informant in a classroom rather than actual words native to TS. For example, 栗 is given the reading [løt^{21}] in the character reading data (c.f. GZ [lœt^{22}]), but the actual word for chestnut is given as 栗子 [lut^{21} tə]. Since the latter is an actual word and the former is likely not, we may safely omit it. The same is true of the finals [-øŋ] and [-øk], as we discussed above. They only occur in character readings, but never in the lexical study.

Zhān, Cheung, et al. (1987a,b)'s EP transcription contains a few "hapax finals." Many do not make it into their initial listing of segments. More confusingly, these transcription idiosyncrasies are not consistent between the character reading and lexicon transcriptions. In the characters, the final [-it] only occurs once in the reading for 月. In the lexicon, 月 is written with the widely attested final [-iet], but now 結 is transcribed once with the hapax final [-iət] in the phrase 結冰 "form ice" but again "normally" as [-iet] when used as a loan graph for the word [kiət^{55}] "hold under the armpit" (Zhān, Cheung, et al., 1987a, pp.8, 197). Similar cases occur for bilabial stop coda finals. [-iep] is a hapax final in the character reading for 捏. [-ip] occurs sporadically with dominant [-iəp]. No similar distinction is found in the finals ending in the bilabial nasal. We judge these variants to all be instances of /-iC/ that were not regularized during the editing process. Like with TS, we preserve the original notation, including non-phonemic [e] or [ə] before /i/ in parentheses.

p/p	p’/p^{h}	^{m}b/m	f/f	v/v	
t/t		^{n}d/n		l/l	
ts/ts	ts’/ts^{h}		s/s		
		^{ŋ}gi/ŋ^{j}		z(i)/z	
k/k	k’/k^{h}	^{ŋ}g/ŋ	h/h		
ku/k^{w}	k’u/k^{wh}				
a/a	^{u}ɔ/ɔ	i/i	u/u	ia/æ	iɔ/iɔ
ai/aj		ei/ij	ui/uj	æi/æj	
au/aw		iu/iw	ou/uw	æu/æw	
am/am		im/im		æm/æm	
an		in/in	un/un	æn/æn	
aŋ/aŋ	ɔŋ/ɔŋ	eŋ/iŋ	ouŋ/uŋ	iaŋ/æŋ	iɔŋ/iɔŋ
ap/ap		ip/ip		æp/æp	
at/at		it/it	ut/ut	æt/æt	
ak/ak	ɔk/ɔk	ek/ik	ouk/uk	iak,iat/æk	iɔk/iɔk
ŋ̍/ŋ̍					
23/1	45/3		55/7	33/9	
22/2	21/4	31/6	21/8		

Table 4.11: XH Transcription Comparison

p/p	p’/p^{h}	^{m}b/m	f/f	v/v	
t/t	t’/t^{h}	^{n}d/n	ɬ/ɬ	l/l	
ts/ts	ts’/ts^{h}		s/s	z/z	
k/k	k’/k^{h}	^{ŋ}g/ŋ	h/h		
a/a	^{u}ɔ/ɔ		i/i	u/u	iɛ/ia
ai/aj	^{u}ɔi/ɔj		ei/ij	ui/uj	
au/aw		eu/ew	iu/iw		iau/iaw
am/am		em/em	im/im		iam/iam
an/an	^{u}ɔn/ɔn	en/en	in/in	un/un	
aŋ/aŋ	ɔŋ/ɔŋ			oŋ,øŋ/uŋ	iaŋ/iaŋ
ap/ap		ep/ep	ip/ip		iap/iap
at/at	^{u}ɔt/ɔt	et/et	it/it	ut/ut	
ak/ak	ɔk/ɔk			oŋ,øk/uk	iak/iak
ŋ̍/ŋ̍					
33/1	55/3		55/7	33/9	
22/2	21/4	31/6	21/8		

Table 4.12: TS Transcription Comparison

p/p	p’/p^{h}	^{m}b/m	f/f	v/v	
t/t	t’/t^{h}	^{n}d/n	ɬ/ɬ	l/l	
ts/ts	ts’/ts^{h}		s/s	z(i)/z	
k/k	k’/k^{h}	^{ŋ}g/ŋ	h/h		
a/a	ɔ/ɔ	i/i	u/u	[p]ɛ,ia/ia	ua/ua
ai/aj	^{u}ɔi/ɔj	ei/ij	ui/uj		uai/uaj
au/aw		iu/iw		ɛu/iaw	
am/am		im/im		ɛm/iam	
an/an		in/in	un/un	(ɛn/ian)	uan/uan
aŋ/aŋ	ɔŋ/ɔŋ	en/iŋ	oŋ/uŋ	iaŋ/iaŋ	
ap/ap		ip/ip		iap/iap	
at/at		it/it	ut/ut	(ɛt/iat)	uat/uat
ak/ak	ɔk/ɔk	et/ik	ok/uk	iak/iak	
m̩/m̩					
33/1	55/3		55/7	33/9	
22/2	21/4	31/6	21/8		

Table 4.13: KP Transcription Comparison

p/p	p’/p^{h}	^{m}b/m	f/f	v/v	
		^{n}d/n		l/l	
ts/ts	ts’/ts^{h}		s/s	z(i)/z	
k/k	k’/k^{h}	^{ŋ}g/ŋ	h/h		
a/a	i/i	u,ɥ/u	ia/ia	ua/ua	(ɔ/ɔ)
ai/aj	ei,([p]ɔi)/ij	ui/uj		uai/uaj	
au/aw	iu/iw		iau/iaw		
am/am	iəm,im/i(ə)m		iam/iam		
an/an	ien/in	un/un	ian/ian	ua/uan	
aŋ/aŋ	eŋ/iŋ	oŋ/uŋ	iaŋ/iaŋ	ɔŋ/uaŋ	iɔŋ/iɔŋ
ap/ap	iəp/ip		iap/iap		
at/at	iet/it	ut/ut	iat/iat	uat/uat	
ak/ak	ek/ik	ok/uk	iak/iak	ɔk/uak	iɔk/iɔk
m̩/m̩					
33/1	55/3		55/7	33/9	
22/2		31/6	21/8		

Table 4.14: EP Transcription Comparison

4.6 The Yōngxún Group

Xiè (2007)'s transcriptions are less consistent than the Zhān and Cheung et al. surveys. In general, he is less concerned with phonological analysis. Many of his finals only occur in one or two readings of low frequency characters and some of his initials are in obvious complementary distribution. For example, in dialects with the the fricativized apical vowel [ɿ], which we write as /z̩/, he often used alveolar sibilants in his transcription despite them not contrasting with his set of post-alveolar sibilants. Fortunately, the dialects we use from his survey have relatively few cases of this overspecification. The most complicated situation is his transcription of MS. We will cover this in the "Unclassified Dialects" section.

We write cases of aberrant finals in near-complementary distribution with their usual forms in parentheses as simple emendations inline to preserve the original forms as we did with the EP dialect. Cases of finals that are distinctive and accord with the rest of phonology but are so marginal that they only occur in one or two character readings are put completely in parentheses.

Xiè's transcription has two idiosyncrasies pertaining to the use of medials. First, he always writes out what he believes are medial elements as [i], [u], and [y] and never interprets them as initials. We rewrite [i] and [u] as [j-] and [w-] when they occur after his zero initials. [y] is always part of complex vowel nuclei and not an independent segment when it occurs as one of Xiè's medials.

Second, he lists glottal stop as a distinctive segment, but it only ever occurs before finals where the main vowel is [i] or [u]. To eliminate this suspicious contrast with his zero initial, we rewrite the cases where [i] and [u] are the main vowel of a final and co-occur with a zero initial as /ji/ and /wu/ respectively. Thus, the contrast between [ik] : [ʔik], for example, is preserved as /jik/ : /ik/ and [ui] : [ʔui] as /wuj/ : /uj/.

p/p	p’/p^{h}	m/m	f/f				
t/t	t’/t^{h}	n/n	ɬ/ɬ	l/l			
tʃ/tʃ	tʃ’/tʃ^{h}	ɲ/ɲ	ʃ/ʃ	i/j			
k/k	k’/k^{h}	ŋ/ŋ	h/h	∅,ʔ/∅			
ku/k^{w}	k’u/k^{wh}			u/w			
a/a		ɛ/ɛ	œ/o	ɔ/ɔ	i/i	u/u	y/y
ai/aj	ɐi/ɐj		oj/oj			ui/uj	
au/aw	ɐu/ɐw	iɛu,iau/(i)ɛw			iu/iw		
an/an	ɐn/ɐn	ɛn/ɛn	on/on		in/in	un/un	yn/yn
aŋ/aŋ	ɐŋ/ɐŋ	iaŋ/ɛŋ	oŋ/oŋ	uaŋ/ɔŋ	iŋ/iŋ	uŋ/uŋ	
am/am	ɐm/ɐm	ɛm/ɛm	om/om		im/im		
at/at	ɐt/ɐt	ɛt/ɛt	œt,ot/ot		it/it	ut/ut	yt/yt
ak/ak	ɐk/ɐk	iak/ɛk	ok/ok	uak/ɔk	ik/ik	uk/uk	
ap/ap	ɐp/ɐp	ɛp/ɛp	op/op		ip/ip		
n̩/n̩							
55/1	35/3	53/5	55/7	33/9			
32/2	23/4	21/6	22/8				

Table 4.15: BL Transcription Comparison

p/p	p’/p^{h}	m/m	f/f			
t/t	t’/t^{h}	n/n	ɬ/ɬ	l/l		
tʃ/tʃ	tʃ’/tʃ^{h}	ɲ/ɲ	ʃ/ʃ	i/j		
k/k	k’/k^{h}	ŋ/ŋ	h/h	∅,ʔ/∅		
ku/k^{w}	k’u/k^{wh}			u/w		
a/a		ɛ/ɛ	ɔ/ɔ	i/i	u/u	(y/y)
ai/aj	ɐi/ɐj		ɔi/ɔj		ui/uj	
au/aw	ɐu/ɐw	iɛu/ɛw		iu/iw		
an/an	ɐn/ɐn	iɛn,ɛn/(i)ɛn	ɔn/ɔn	in/in	un/un	
aŋ/aŋ	ɐŋ/ɐŋ	iɛŋ/ɛŋ	ɔŋ/ɔŋ	iŋ/iŋ	uŋ/uŋ	
am/am	ɐm/ɐm	iɛm/ɛm		im/im		
at/at	ɐt/ɐt	iɛt/ɛt	ɔt/ɔt	it/it	ut/ut	
ak/ak	ɐk/ɐk	iɛk,ɛk/(i)ɛk	ɔk/ɔk	ik/ik	uk/uk	
ap/ap	ɐp/ɐp	iɛp/ɛp		ip/ip		
ɻ/ʐ̩	ŋ̍/ŋ̍					
55/1	35/3		33/5	55/7	33/9	
21/2	23/4		22/6	22/8		

Table 4.16: QZ Transcription Comparison

p/p	p’/p^{h}	m/m	f/f		
t/t	t’/t^{h}	n/n	ɬ/ɬ	l/l	
tʃ/tʃ	tʃ’/tʃ^{h}	ɲ/ɲ	ʃ/ʃ	i/j	
k/k	k’/k^{h}	ŋ/ŋ	h/h	∅,ʔ/∅	
ku/k^{w}	k’u/k^{wh}			u/w	
a/a		ɛ/ɛ	ɔ/ɔ	i/i	u/u
ai/aj	ɐi/ɐj		ɔj/ɔj		ui/uj
au/aw	ɐu/ɐw	ɛu/ɛw		iu/iw	
an/an	ɐn/ɐn	ɛn/ɛn	ɔn/ɔn	in/in	un/un
aŋ/aŋ	ɐŋ/ɐŋ	ɛŋ/ɛŋ		iŋ/iŋ	uŋ/uŋ
am/am	ɐm/ɐm	ɛm/ɛm		im/im	
at/at	ɐt/ɐt	ɛt/ɛt	(ɔt/ɔt)	it/it	ut/ut
ak/ak	ɐk/ɐk	ɛk/ɛk	(ɔk/ɔk)	ik/ik	uk/uk
ap/ap	ɐp/ɐp	ɛp/ɛp		ip/ip	
53/1	33/3	45/5	55/7	33/9	
31/2	23/4	22/6	22/8	23/10	

Table 4.17: NN_p Transcription Comparison

p/p	p’/p^{h}	m/m	f/f				
t/t	t’/t^{h}	n/n	ɬ/ɬ	l/l			
tʃ/tʃ	tʃ’/tʃ^{h}	ɳ/ɲ	ʃ/ʃ	i/j			
k/k	k’/k^{h}	ŋ/ŋ	h/h	∅,ʔ/∅			
ku/k^{w}	k’u/k^{wh}			u/w			
a/a		ɛ/ɛ	ɯ/ə	ɔ/ɔ	i/i	u/u	y/y
ai/aj	ɐi/ɐj			ɔi/ɔj		ui/uj	
au/aw	ɐu/ɐw	ɛu/ɛw			iu/iw		
an/an	ɐn/ɐn	ɛn/ɛn	ən/ən	ɔn/ɔn	in/in	un/un	
aŋ/aŋ	ɐŋ/ɐŋ	ɛŋ/ɛŋ	əŋ/əŋ	ɔŋ/ɔŋ	iŋ/iŋ	uŋ/uŋ	
am/am	ɐm/ɐm	ɛm/ɛm			im/im		
at/at	ɐt/ɐt		ət/ət	ɔt/ɔt	it/it	ut/ut	
ak/ak	ɐk/ɐk	ɛk/ɛk		ɔk/ɔk	(ik/ik)	uk/uk	
ap/ap	ɐp/ɐp	ɛp/ɛp			ip/ip		
53/1	33/3	45/5	55/7	33/9			
31/2	23/4	22/6	22/8	23/10			

Table 4.18: BS_p Transcription Comparison

p/p	p’/p^{h}	m/m	f/f				
t/t	t’/t^{h}	n/n	ɬ/ɬ	l/l			
tʃ/tʃ	tʃ’/tʃ^{h}	ɳ/ɲ	ʃ/ʃ	i/j			
k/k	k’/k^{h}	ŋ/ŋ	h/h	∅,ʔ/∅			
ku/k^{w}	k’u/k^{wh}			u/w			
a/a		ɛ/ɛ	œ/œ		i/i	əu,u/u	y/y
ai/aj	ɐi/ɐj			əi/əj		ui/uj	
au/aw	ɐu/ɐw	ɛu/ɛw	œu/œw		iu/iw		
an/an	ɐn/ɐn	ɛn/ɛn	œn/œn	ən/ən	in/in	un/un	yn/yn
aŋ/aŋ	ɐŋ/ɐŋ	ɛŋ/ɛŋ	œŋ/œŋ	əŋ/əŋ	iŋ/iŋ	uŋ/uŋ	
am/am	ɐm/ɐm	ɛm/ɛm	œm/œm		im/im		
at/at	ɐt/ɐt	ɛt/ɛt	œt/œt	ət/ət	it/it	ut/ut	yt/yt
ak/ak	ɐk/ɐk	ɛk/ɛk	œk/œk	ək/ək	ik/ik	uk/uk	
ap/ap	ɐp/ɐp	ɛp/ɛp	œp/œp		ip/ip		
35/1	33/3	55/5	55/7	33/9			
213/2	22/4	52/6	52/8	22/10			

Table 4.19: BY Transcription Comparison

4.7 Unclassified Dialects

In MS, the transcription of the finals [-ɔi] and [-uɔi],[-ɔn] and [-uɔn], [-uəŋ] and [-uɔŋ], and [-uək] and [-uək] never bear distinction. We merge them all with the main vowel /ɔ/ without creating any minimal pairs. However, conditioning environments for the surface allophones of these seems ad hoc in many cases. For example, [-uɔi] only ever occurs twice. Once in the character reading for 該 [kuɔi^{53}] and again in the character reading for 率 [ʃuɔi^{44}]. These initials never occur with [-ɔi]. [-ɔn] only occurs after the zero initial and [-uɔn] is used elsewhere. [-uəŋ] and [-uək] are used after velar initials and [f-], [-uɔŋ] and [-uɔk] elsewhere. For his glottal stop initial, Xiè writes [ʔuɔk] and [ʔuəŋ], but has no contrasting [ʔuək] and [ʔuɔŋ].

His nucleus [ə] is in complementary distribution with the phonemic nucleus /ɔ/ [uə ũɔ ɔ̃] we discussed above. It only occurs after acute initials in the two distinctive finals [-ən] and [-ət], while [-(u)ɔn] and [-(u)ɔt] only occur after graves. We do not write these as one phoneme for clarity and since this reduction seems questionable. If these two are allophones, they do not share many features. Would a conditioning environment of acute onsets and codas around this nucleus really cause unrounding and centralization?

WC [ɛi] and [ɔu] are in complementary distribution with /i/ and /u/ respectively. [i] occurs after palatal initials, [ɛi] elsewhere. [u] occurs after velars, /f/ and /w/, [ɔu] elsewhere.

p/p	p’/p^{h}	m/m	f/f				
t/t	t’/t^{h}	n/n	ɬ/ɬ	l/l			
tʃ/tʃ	tʃ’/tʃ^{h}	ɲ/ɲ	ʃ/ʃ	i/j			
k/k	k’/k^{h}	ŋ/ŋ	h/h	∅,ʔ/∅			
ku/k^{w}	k’u/k^{wh}			u/w			
a/a		ɛ/ɛ		ɔ/ɔ	i/i	u/u	ua/ua
ai/aj	ɐi/ɐj			uɔi,ɔi/(u)ɔj		ui/uj	uai/uaj
au/aw	ɐu/ɐw	iɛu,iau/(i)ɛw	əu/əw		iu/iw		
an/an	ɐn/ɐn	ɛn/ɛn	ən/ən	ɔn,uɔn/ɔn	in/in	un/un	uan/uan
aŋ/aŋ	ɐŋ/ɐŋ	iaŋ/ɛŋ		uəŋ,uɔŋ/ɔŋ	iŋ/iŋ	uŋ/uŋ	uaŋ/uaŋ
am/am	ɐm/ɐm	iam/ɛm	əm/əm		im/im		
at/at	ɐt/ɐt	ɛt/ɛt	ət/ət	uɔt/ɔt	it/it	ut/ut	uat/uat
ak/ak	ɐk/ɐk	iak/ɛk		uək,uɔk/ɔk	ik/ik	uk/uk	uak/uak
ap/ap	ɐp/ɐp	iap/ɛp		əp/ɔp	ip/ip		
53/1	42/3	44/5		55/7	33/9		
21/2	35/4	214/6		22/8			

Table 4.20: MS Transcription Comparison

p/p	p^{h}/p^{h}	m/m	f/f		
t/t	t^{h}/t^{h}	n/n	ɬ/ɬ	l/l	
tʃ/tʃ	tʃ^{h}/tʃ^{h}		ʃ/ʃ	j/j	
k/k	k^{h}/k^{h}	ŋ/ŋ	h/h		
				w/w	
a/a		ɔ/ɔ	i/i	u/u	ɛ/iɛ
ai/aj	ɐi/ɐj	ɔi/ɔj	ei/ij	ui/uj	
au/aw	ɐu/ɐw	ou/ɔw	iu/iw		iɛu/iɛw
an/an	ɐn/ɐn	ɔn/ɔn	in/in	un/un	iɛn/iɛn
aŋ/aŋ	ɐŋ/ɐŋ	ɔŋ/ɔŋ	eŋ/iŋ	uŋ/uŋ	
am/am	ɐm/ɐm		im/im		iɛm/iɛm
at/at	ɐt/ɐt	ɔt/ɔt	it/it	ut/ut	iɛt/iɛt
ak/ak	ɐk/ɐk	ɔk/ɔk	ek/ik	uk/uk	
ap/ap	ɐp/ɐp		ip/ip		iɛp/iɛp
33/1	21/3	24/5	24/7	21/9	
42/2		55/6	55/8	42/10	

Table 4.21: YJ Transcription Comparison

	p’/p^{h}	b/b	m/m	f/f	
t/t	t’/t^{h}	d/d	n/n		l/l
tʃ/tʃ	tʃ’/tʃ^{h}		ɲ/ɲ	ʃ/ʃ	j/j
k/k	k’/k^{h}		ŋ/ŋ	h/h	
kw/k^{w}	k’w/k^{wh}		ŋw/ŋ^{w}		w/w
a/a		ɛ/ɛ	ɔ/ɔ	i,ɛi/i	u,ɔu/u
ai/aj	ɐi/ɐj		ɔi/ɔj		ui/uj
au/aw	ɐu/aw	ɛu/ɛw	ɔɒu/ɔw	iu/iw	
am/am	ɐm/ɐm	ɛm/ɛm		im/im	
an/an	ɐn/ɐn	ɛn/ɛn	ɔn/ɔn	in/in	un/un
aŋ/aŋ	ɐŋ/ɐŋ	ɛŋ/ɛŋ	ɔŋ/ɔŋ	iŋ/iŋ	uŋ/uŋ
ap/ap	ɐp/ɐp	ɛp/ɛp		ip/ip	
at/at	ɐt/ɐt	ɛt/ɛt	ɔt/ɔt	it/it	ut/ut
ak/ak	ɐk/ɐk	ɛk/ɛk	ɔk/ɔk	ik/ik	uk/uk
55/1	13/3	21/5	4/7	2/9	
44/2	33/4	22/6	3/8		

Table 4.22: WC Transcription Comparison

Chapter 5

TONES

Yue-Hashimoto (2000) reconstructs nine tones in her proto-Yue. We largely accept her correspondences but add two revisions. First, we eliminate the *zhōngrù* 中入 tone, the tone 9 of our correspondences, from her system as a toneme. Tone 9, though found in almost all modern Yuè dialects, cannot be projected back to the earliest common ancestor because CY complex vowel nuclei, /ia/ and /iɔ/, seem to have monophthonginzed to mid vowels that did not condition tone lowering before codas involving closure at the front of the oral cavity in the Sìyì dialects. Furthermore, the MS dialect and some dialects in Guǎngxī, such as the Yùlín 鬱林 dialect, do not distinguish between tones 7 and 9 (Tsuji, 1980). The NN_p and BS_p dialects have tone 9 for all cases of CY tone 7 except in etymon #015 BIRD (2) 雀 /*ts^{h}iɔk^{7}/ which bears tone 7. This word also exhibits irregular aspiration in the onset in these dialects, so these dialects' reflexes of this etymon may not be the result of regular sound change. We view the development of tone 9 as a wave of sound change that originated in the Pearl River Delta region and radiated outward along the surrounding river ways due to its seemingly incomplete nature in the Sìyì and inland dialects. The conditioning environment for tone 9, i.e., vowel nuclei that ended in open vowels, had already disappeared in some environments before the wave reached the Sìyì

region, preventing further tone change.

Second, we give new correspondences for tones 3 and 5 when they occur with pre-glottalized sonorant onsets. These correspondences have not been covered in any previous study of Yuè historical phonology. These onsets correspond to similar voiceless sonorants in Common Min and Early Southern Highlands Chinese, but in those systems they are usually reconstructed with lower register tones. Coblin (n.d.[a]) follows Norman and treats the high register reflex as a development conditioned by voicelessness, ultimately going back to a low register tone in the common system. This development seems unlikely. If the registers had already formed and these initials were in the low register, how would voicelessness raise them to the high register? We regard the entire formation of high *yīn* register tones as conditioned by the features of onsets. In the case of these sonorant onsets, we have chosen pre-glottalized as the relevant feature on the presumption that the glottalization would have entailed delayed voice onset time phonetically, similar to the plain and aspirated obstruents and fricatives. The voiced obstruent and plain sonorants would have carried over voicing from previous utterances, giving them a negative voice onset time that in turn lowered tone register. We justify our choice of pre-glottalized sonorants rather than voiceless sonorants in their relevant section in the following chapter.

Table 5.1 demonstrates the basic tone correspondences. The XH dialect occasionally demonstrates fluctuation of tone 7 and 9 reflexes in syllables that descend from reflexes of CY complex nuclei. This is likely due to dialect mixture resulting in convergence towards the Guǎngfǔ trend. The XH dialect shares the closest border with the Guǎngfǔ group of

the Sìyì dialects discussed here. The YS and HJ dialects exhibit both upper and lower register reflexes of tone 3 in words that began with pre-glottalized sonorants in CY. This is probably also due to convergence. A few cases of tone 4 with sonorant initials in the YS dialect also take tone 3. The DG dialect features many more tone splits of [+open] vocalic nuclei in stop coda syllables. These splits are conditioned by the coda itself along with other vowel features. Convergence has partially interrupted these splits, creating a very complicated correspondence situation. We discuss this development in further detail in the nucleus chapter.

Tone Cond.	1	2	3 obs.	3 son.	4	5 obs.	5 son.	6	7 VV{t,p}	7 open	7	8	8 open
YS	1	2	3	3/4	3/4	5	5	6	9	9	7	8	8
LX	1	2	3	4	4	5	5	6	8/9	8/9	7	8	8
HJ	1	2	3	3/4	4	5	5	6	9	9	7	8	8
FK	1	2	3	4	4	5	5	6	9	9	7	8	8
GZ	1	2	3	4	4	5	5	6	9	9	7	8	8
NH	1	2	3	4	4	5	5	6	9	9	7	8	8
DG	1	2	3	4	4	5	5	5	9/5	9/5	7	8	8
FG	1	2	3	3	2	1	1	6	9	9	7	8	8
YF	1	2	3	4	6	5	2	2	9	9	7	8	8
XH	1	2	3	3	1/4	1	1	6	9/7	9	7	8	8
TS	1	2	3	3	1/4	1	1	6	7	9	7	8	8
EP	1	2	3	3	1/6	1	1	6	7	9	7	8	8
KP	1	2	3	3	1/4	1	1	6	7	9	7	8	8
BL	1	2	3	4	4	5	5	6	9	9	7	8	8
QZ	1	2	3	4	4	5	5	6	9	9	7	8	8
NN_p	1	2	3	4	4	5	5	6	9	9	9	8	10
BS_p	1	2	3	4	4	5	5	6	9	9	9	8	8
BY	1	2	3	4	4	5	5	6	9	9	9	8	10
MS	1	2	3	4	4	5	5	6	7	7	7	8	8
YJ	1	2	3	3	3	5	5	6	9	9	7	8	8
WC	1	2	3	4	4	5	5	6	9	9	7	8	8

Table 5.1: Common Yue Tone Correspondence

5.1 The Development of Tone 9

Departing from previous explanations, CY treats the split of tone 7 into its upper 7 and lower 9 variants as conditioned by vowel openness. Previous reconstructions of Yuè historical phonology treat this split as conditioned by a tense/lax or length distinction in vowels (McCoy, 1966; Tsuji, 1980; Yue-Hashimoto, 2002). These explanations are un-

satisfactory for two reasons. 1) The tense/lax distinction is disputed as a phonological feature and does not correlate with any one phonetic feature that would have probably conditioned tone change. Vowel height, on the other hand, has been found to correlate with intrinsic lower F0. Ohala et al. (1979, p.52) proposed that this phonetic lowering presents an excellent conditioning environment for tone lowering. This theoretical proposal matches the situation reflected in Common Yue. 2) Tenseness and length are not minimally contrastive phonological features in any modern Yuè dialect. The contrast between /ɐ/ and /a/ may easily be accounted for by a phonemic difference in openness, i.e., vowel height, rather than tenseness. This is preferable because it also accounts for the featural distinction between the pairs i/ɛ, y/œ, u/ɔ, etc. with a minimal set of contrastive features. Furthermore, the vowel systems of many other Chinese dialects may also be accounted for with the same minimal set of contrastive features [open], [front], and [round]. A modern subphonemic status does not necessarily disqualify a tense/lax distinction as a conditioning environment, but one must explain the basis for a historical tense/lax distinction in terms separate from the development of tone 9 itself. Using the development of tone 9 as the sum and substance of the argument for a historical tense/lax distinction is circular reasoning. Common Yue's open vowels are supported by modern dialect reflexes and universal mechanics of vowel chain shifts rather than being tied solely to tone 9.

Yue-Hashimoto (2002) accounts for the development of tone 9 by reconstructing all modern Pearl River Delta tone 9 words with long vowels. She tentatively reconstructs the

finals given in table 5.2. I have supplied Common Yue nuclei to facilitate comparison. Yue-Hashimoto does not reconstruct a final to account for the /-ɐp/ final of Common Yue. Yue-Hashimoto gives no examples of her /yɔː/ nucleus, so I cannot relate it to anything in Common Yue. Yue-Hashimoto's distinction of the nuclei /ɑː/ and /ɔː/ is marginal and only occurs before velar codas in particular character readings. It seems like the latter /ɔː/ corresponds to /uɔ/ in our system, but she does not give specific examples of this final, so it is difficult to be sure.

CY Nucleus	Yue-Hashimoto Nucleus	Possible Checked Codas
i	i	-k
(w)i	ui	-k
u	o	-k
ɐ	a	-k, -t
ɐ	?	-p
i	a	-t
i	ia	-p
y	ya	-t
u	ua	-t
a	aː	-k, -t, -p
ɔ	ɑː	-k, -t, -p
uo	ɔː	-k
ia	iaː	-k, -t
ia	iɛː	-p
(w)a	uaː	-k, -t
iɔ	iɔː	-k
iɔ	yaː	-t
(w)ɔ	uɔː	-k
uɔ	uɑː	-t
?	yɔː	-k

Table 5.2: A Comparison of Common Yue Vowel Nuclei and Yue-Hashimoto's Tentative Proto-Yue Vowel Nuclei

Her system is unnatural. There are more diphthongs than monophthongs and a large portion of the vocalic nuclei are in complementary distribution. Her monophthongs suffer from an unrealistically restricted distribution. The length distinction does not correlate with any modern phonological contrast other than the presence of tone 9. It is mechanical rewriting, not an explanation of tone 9's development.

Huang (2009) argues that the development of tone 9 was conditioned by the feature advanced tongue root [ATR]. Syllables with nuclei marked [-ATR] took the high register, while [+ATR] counterparts took the low register. She reasons that because [ATR] is "usually accompanied by lowering of the larynx *or* a slightly breathy quality which lowers the tone" (emphasis supplied) (Huang, 2009, p.7). This statement is confusing. Which lowered the tone in the case of proto-Yue? Was it conditioned by an intermediate, subphonemic stage of lowered larynx *or* breathy voice? The two are related but not synonymous. Regardless, the explanation is unsatisfactory. No modern Yuè dialect reflects phonetic traces of larynx lowering or breathy voice corresponding to her [+ATR] vowels. [ATR] is also not otherwise necessary to account for the phonemic contrasts of vowel nuclei diacrhonically in Yuè dialects, nor is it a minimally contrastive feature in any modern Yuè dialect. In the end, it is yet another mechanical rewriting of tone 9 rather than an explanation for the development that stands alone. Moreover, Huang's vowel inventory far out numbers any modern Yuè dialect. She even has redundant proto-forms, /*i̯u, *y,/ that share identical correspondences. Within her paper, she reconstructs "snow" 雪 twice as /*syt/ first, then again as /*si̯ut/ (Huang, 2009, pp.8, 16). Like Yue-Hashimoto (2002),

many of her vowel nuclei appear to be in complementary distribution.

The explanation CY offers is based on conservative reflexes of vowel nuclei found preserved in dialects of the Fēnglián and Sìyì groups. For example, in the case of the Sìyì group, high vowels in checked syllables in TS typically predict the retention of tone 7 in other dialects, with the notable exception of CY /*iɔ/ which raised to /y/ in the Sìyì dialects. TS mid /e/ developed from CY /*ia/, and TS /ia/ developed from CY /*ia/ and /*iɔ/. These vowel nuclei reliably correlate with the development of tone 9 elsewhere. These sorts of conservative reflexes occur in other Sìyì dialects as well, but are partially obscured by other innovations.

In the LX dialect, CY /*ia/ is retained as open /ɛ/ before dentals codas and after dental onsets rather than raising as it did elsewhere. As for the Guǎngfǔ group, in the eastern Zēngchéng 增城 dialect, not included in our correspondences, the finals /ɛŋ/ and /ɛk/ correspond with CY /*ian/ and /*iat/, and the finals /œŋ/ and /œk/ correspond with /*iɔn/ and /*iɔt/. This pattern is occasionally interrupted by the usual Guǎngfǔ reflexes /in, it/ and /yn, yt/, respectively, as the result of convergence.

For the dialects of LX and Bùtián of the northern inland region, the Zēngchéng dialect in the east, and TS dialect in the south to all preserve nonclose vowels in these finals cannot be ascribed to common innovation. They are too geographically disparate. This must have been the original situation, now only preserved on the fringes of the greater dialect region.

Our discussion of CY vowel nuclei will detail the conditions in which original [+open]

vowels rose in modern Yuè dialects, masking the original conditioning environment of the split. This was part of a chain shift of vowels that affected most of the region.

A similar split also occurred in tone 8 in the NN_p and BY dialects. Xiè (2007) noted what I label as tone 10 in different ways in these two dialects. He has an upper *shàng yángrù* 上陽入 and lower *xià yángrù* 下陽入 tone for both dialects, but applies the labels by their relative phonetic height in each dialect rather than labeling them as a phonemic distinction that crosses dialect boundaries. His lower *xià yángrù* tone of the BY dialect actually corresponds to his upper *shàng yángrù* tone of the NN_p dialect. I relabeled these so that the BY tones are switched so the two line up appropriately diaphonemically as tones 8 and 10 respectively. The lowering condition is the same as it is for the split of tone 7 into tones 7 and 9.

5.2 Upper Register Tones and Pre-glottalized Sonorant Onsets

The correspondences for tones 1, 5, and 7 coupled with CY pre-glottalized sonorant onsets are the same as for the correspondences of these tones coupled with obstruent onsets. Currently, we only have one proper example of a pre-glottalized sonorant onset occurring with tone 5 in CY, #188 Pretty 靚 /*ʔliaŋ^{5}/. Pre-glottalized sonorant onsets coupled with open vowel nuclei in checked syllables take tone 9, similar to the development of tone 9 in other voiceless environments. The correspondences for tone 3 coupled with pre-glottalized sonorant onsets differ from the obstruent correspondences. In the case of tone 3, only the Sìyì group preserves the high register reflexes consistently.

5.3 Tone 4 in the Sìyì Group

In some words, tone 4 corresponds to tone 1 in the Sìyì group. Even though the Sìyì group has an independent tone 4, this seems to be the dominant trend. Tone 4 in the Sìyì group does double duty as a morphologically "changed tone" that marks a diminutive variant of words. Because of this, it is difficult to tell if the cases of Sìyì tone 4 corresponding to tone 4 elsewhere are a case of retention or innovation. When Sìyì tone 4 corresponds to tone 4 elsewhere, it occurs most often with sonorant onsets, except in a few etyma, like #074 MATERNAL UNCLE 舅 /*giw^{4b}/. These may ultimately be cases of morphological "changed tone" in the Sìyì dialects since the low falling changed tone is the same in surface form as lexical tone 4. If we accept this interpretation, then it seems to be the case that the Sìyì group had already mostly merged tone 4 with tone 1 in a stage of the language before the introduction of *koiné* lexicon. In the *koiné*, tone 4 only survived before sonorant onsets. With obstruent onsets, it had merged with tone 6. This is why Common Yue tone 4, when it corresponds to Sìyì group tone 4, usually occurs with sonorant onsets. It reflects the Sìyì group attempting to accommodate northern tone 4 after the original tone 4 had already merged with tone 1. There are still colloquial cases of tone 1 merged tone 4 in the case of sonorant onsets in the Sìyì group, for example the second person pronoun #180 2SG /*ni^{4a}/.

Yue-Hashimoto (2000; 2005) rejected the correspondence of tone 1 in the Sìyì group with tone 4 elsewhere because it appeared to be influenced by Hakka dialects. We do not

reject the correspondence so as not to bias the reconstruction since it is regular and widely attested. However, Yue-Hashimoto is likely correct that there has been outside influence. The evolution of tone 4 to tone 1 is a characteristic feature of Hakka dialects used as a criterion to discriminate between them and other surrounding dialects (Norman, 1988, p.333).

Yue-Hashimoto (2005) argues that the development of tone 4 into tone 1 is a late innovation in TS dialect due to contact with Hakka dialects. Kwok (2006, p.711) joins her in this view, but expands the contact from just the TS dialect to all Sìyì group dialects. This of course implies a greater time depth since all dialects of the group share the supposed innovation, as we can see clearly in Common Yue correspondences.

Yue-Hashimoto (2005) ultimately proposed that TS "borrowing" the change of tone 4 to tone 1 and TS sharing other characteristic common innnovations with other Yuè dialects implies that the formation of Gàn and Hakka postdates the formation of Yuè as a dialect family. We do not adopt this stance. Coblin (2015) demonstrated that Gàn dialects share substrata with Hakka dialects, but the formation of Common Gan cannot be traced before contact with northern *koiné*, so the Gàn part of the equation is irrelevant. More interestingly, the merger of tone 1 and tone 5, characteristic of the whole Sìyì group, is also a defining characteristic of Shē 畲 dialects (Nakanishi, 2010). This merger in Shē postdates the split of Shē dialects from Common Hakka-Shē (Coblin, n.d.[b]). Working out the exact relationship between the Sìyì group and Common Shē would greatly inform our current understanding of the chronology of the formation of Common Yue and Common

Hakka-Shē and the relationship between the two.

For now, it seems likely that the early Yùe linguistic communities of the Sìyì region had contact with Shē speakers. Beyond just the merger of tone 1 with tones 4 and 5, many Sìyì dialects have reflexes of the first person prounoun with a final /-j/ coda similar to how this pronoun is realized in Shē dialects (Nakanishi, 2010, p.248). Today, Shē speakers are found along the eastern coast of Guǎngdōng, but it is not unreasonable to assume that they may have also originally settled slightly west of the Pearl River Delta as well in the Sìyì region. In CY, we will mark instances of tone 4 corresponding to Sìyì tone 1 as tone 4a corresponding to Sìyì tone 4 as 4b to make the differing reflexes explicit. It is not yet clear which one of these reflexes of tone 4 is older. It could be the case that for tone 4b some reflexes go back to a stage before contact with Shē, while others owe to later literary readings of characters with sonorant onsets.

5.4 Examples of the Tones

Table 5.3 provides examples of etyma for each of the CY tones. The etyma are indexed by semantic domain and English gloss in alphabetical order. Characters corresponding to the words are written in commonly accepted orthography. Cases of loan graphs or "vulgar" characters (*súzì* 俗字) are underlined. I put characters in parentheses that write words in literary Sinitic that might have distant etymological links with CY words. A full listing of the data for each etymon is provided in Appendix C.

Tone	Examples			
1	#192	BODY	身	/*ʃin^{1}/
	#206	FACIAL HAIR	鬚	/*su^{1}/
	#195	BREASTs		/*ʔnɛn^{1}/
	#045	MOSQUITO	蚊	/*ʔmun^{1}/
2	#214	HEAD	頭	/*dɐw^{2}/
	#219	LIPS	脣	/*ʒyn^{2}/
	#022	COW	牛	/*ŋɐw^{2}/
3	#222	NECK	頸	/*kiaŋ^{3}/
	#203	EYE	眼	/*ʔŋan^{3}/
	#213	HAND	手	/*ʃiw^{3}/
	#061	FEMALE SUFF.	嫲 (女)	/*ʔna^{3}/
4	#041	HORSE	馬	/*ma^{4b}/
	#185	INSIPID	淡	/*dam^{4a}/
	#074	MATERNAL UNCLE	舅	/*giw^{4b}/
5	#037	HARE	兔	/*tʰu^{5}/
	#054	RIBBONFISH	帶	/*taj^{5}/
6	#205	FACE	面	/*mian^{6}/
	#239	TONGUE	脷	/*li^{6}/
	#200	COWLICK (HAIR)	(旋)	/*dziɔn^{6}/
	#029	EGG (1)	蛋	/*dan^{6}/
7	#232	SPINE	脊骨	/*tsiak^{7} kut^{7}/
	#208	FINGERNAIL	甲	/*kap^{7}/
	#035	FROG	蛤	/kɐp^{7}/
8	#241	WING	翼	/*jik^{8}/
	#052	PIGEON	白鴿	/bak^{8} kɔp^{7}/
	#017	BUTTERFLY	蝴蝶	/*ɣu^{2} diap^{8}/

Table 5.3: Examples of the Common Yue Tones

Chapter 6

ONSETS

CY initials largely follow the general onset inventory of Yue-Hashimoto (2006) with six important changes. First, voiced onsets are not further subdivided into phonetic aspirated and unaspirated subvarieties. Second, we recast the distinction between Yue's retroflex and alveo-palatal series as a distinction of syllable nucleus rather than a distinction of onset. Yue's retroflex initial series may only be reconstructed before [i]-like open-syllable finals. We treat this correspondence as instances of a "super-high" fricativized apical vowel that we write as /ʐ̩/ after post-alveolar onsets. This nucleus only occurs after our alveolar and post-alveolar series. Third, we add a new series of pre-glottalized sonorants. Fourth, we reject all of Yue's complex onsets. Fifth, we separate Yue's labial fricative into a voiced and voiceless pair. Sixth, we add a series of phonemic labiovelars.

Labial	Dental	Alveolar	Palatal	Velar	Lab. Velar
p	t	ts	tʃ	k	k^w
p^h	t^h	ts^h	$tʃ^h$	(k^h)	k^{wh}
b	d	dz	dʒ	g	g^w
f		s	ʃ	x	x^w
v		(z)	ʒ	ɣ	
m	n		ɲ	ŋ	$ŋ^w$
^ʔm	^ʔn		$^ʔɲ$	$^ʔŋ$	$^ʔŋ^w$
	l		j		w
	^ʔl		(^ʔj)		^ʔw

Table 6.1: Common Yue Onset Inventory

6.1 Stops

6.1.1 Plain Stops

For simplicity, we treat affricates as belonging to the same articulatory class as stops at the recommendation of Kehrein (2013). Their development in Yue dialects follows a similar trajectory as other traditional "true stops" so they are best discussed together. Stop onsets in Yuè dialects remain remarkably unchanged. The most noticeable difference is an overwhelming loss of the post-alveolar series. The original situation only survives in the LX dialect in our correspondences, but the Bóbái 博白 and Yīngdé 英德 dialects also partially maintain this distinction between alveolar and post-alveolar sibilants. Though predominantly merged with the alveolar sibilants in the Guǎngfǔ group, consonant onset chain shifts have occurred elsewhere in several dialects that preserve the distinction as a difference of dental stops versus alveolar sibilants. Each of these shifts occurred likely

CY Onset	*p	*t	*ts	*tʃ	*k	*kʷ	pʰ	tʰ	tsʰ	tʃʰ	kʰ	kʷʰ
YS	p	t	ts	ts	k	kʷ	pʰ	tʰ	tsʰ	tsʰ		–
LX	p	t	ts	tʃ	k	kʷ	pʰ	tʰ	tsʰ	tʃʰ		kʷʰ
HJ	p	t	ts	ts	k/ts	kʷ	pʰ	tʰ	tsʰ	tsʰ		–
FK	p	t	ts	ts	k/ts	kʷ	pʰ	tʰ	tsʰ	tsʰ		kʷʰ
GZ	p	t	ts	ts	k	kʷ	pʰ	tʰ	tsʰ	tsʰ		f
NH	p	ʈ/t	t/ts	ts	k	kʷ	pʰ	tʰ	tʰ/tsʰ	tsʰ		w
DG	p	t	ts	ts	k	kʷ	pʰ	tʰ	tsʰ	tsʰ		f
FG	p	t	ts	ts	k	kʷ	pʰ	tʰ	tsʰ	tsʰ		f
YF	p	t	ts	ts	k	kʷ	pʰ	tʰ	tsʰ	tsʰ		f
XH	p	t	ts	ts	k	kʷ	pʰ	h	tsʰ/s	tsʰ/s		f
TS	p	∅	t	ts	k	k	pʰ	h	tʰ/ɬ	tsʰ/ɬ		f
EP	p	t	ts	ts	k	k	pʰ	h	tsʰ/s	tsʰ/s		f
KP	v	∅	t	ts	k	k	pʰ	h	tʰ/ɬ	tsʰ/ɬ		f
BL	p	t	t	tʃ	k	kʷ	pʰ	tʰ	tʰ	tʃʰ		w
QZ	p	t	tʃ	tʃ	k	kʷ	pʰ	tʰ	tʃʰ	tʃʰ		f
NN_p	p	t	tʃ	tʃ	k	kʷ	pʰ	tʰ	tʃʰ	tʃʰ		kʷʰ
BS_p	p	t	tʃ	tʃ	k	kʷ	pʰ	tʰ	tʃʰ	tʃʰ		kʷʰ
BY	p	t	tʃ	tʃ	k/tʃ	kʷ	pʰ	tʰ	tʃʰ	tʃʰ		–
MS	p	t	ɬ	tʃ	k/tʃ	k(u)	pʰ	tʰ	tʰ	tʃʰ		kʰ(u)
YJ	p	t	tʃ	tʃ	k	k	pʰ	tʰ	tʃʰ	tʃʰ		kʰ
WC	b	d	t	tʃ	k	kʷ	pʰ	tʰ	tʰ	tʃʰ		f

Table 6.2: Common Yue Voiceless Stops

independently given the “top” of each chain differs in expression.

- NH tʃ > ts > t > ʈ; tʃ > tsʰ > tʰ (partial merger)
- WC ts > t > d; tsʰ > tʰ (merger)
- TS,KP tʃ > ts > t > ∅; tʃʰ > tsʰ > tʰ > h
- BL ts > t (merger); tsʰ > tʰ (merger)
- MS ts > ɬ (merger with CY /*s/); tsʰ > tʰ (merger)

The chain shift in the NH dialect superficially appears incomplete. Both pre and post-shift reflexes of the CY plain dental and alveolar stops occur. This is due to the accommodation of GZ forms as part of an ongoing process of dialect leveling occurring in the Guǎngfǔ group. We discuss this leveling in more detail below in the voiced stop section.

The WC dialect is rather divergent in its development of laryngeal distinctions in stops. Bilabial and dental plain stops have developed into voiced stops, presumably by way of glottalization and implosivization, similar to Vietic languages. In the KP dialect, the plain bilabial stop has evolved into /v/. This might also be due to a case of intermediate glottalization.

As a general rule for dialects that participate in the greater Yuè dialect vowel chain shift, velars before CY /*u/ in syllables that end in codas that involve stricture at the front of the oral cavity, i.e. end in {-m,-p,-n,-t,-j,-w}, evolve into labiovelars. They retain the previous rounding of /*u/ as the original vowel moved downward along the nonperipheral track towards /ɐ/. For example, CY /*kun/ typically becomes /*kʷɐn/. This widespread change is not covered in Table 6.2, but is accounted for in the correspondences for the vowel nulcei later. Additionally, velars palatalize in the HJ, FK, BY, MS dialects in certain conditions. In MS and BY, the conditioning environment was before CY /*i/ followed by a rounded coda, i.e., /_i{m,p,w}. Palatalization occurs in HJ in the same environments except before the final /-ip/, possibly because the main vowel was lax before a final stop. BY has additional palatalization before the modern nucleus /a/, but this trend is partially disturbed by influences from the surrounding dialects. In FK, the

conditioning environment was before CY /*i/ followed by a coda involving stricture at the front of the oral cavity, i.e., /_i{m,p,n,t,w,j}. Elsewhere, velar stop onsets remained stable.

Original Labiovelars have merged with regular velars in the TS, EP, KP, and YJ dialects. They have remained stable elsewhere.

Onset	Example			
*p-	#076	OLDER PATERNAL UNCLE	伯	/*bak^{7}/
*t-	#54	RIBBONFISH	帶	/*taj^{5}/
*ts-	#077	SON	仔	/*tsɐj^{3}/
*tʃ-	#094	UMBRELLA	遮	/*tʃia^{1}/
*k-	#147	GARLIC CHIVE	韭黃	/*kiw^{3} wɔŋ^{2}/
*k^{w}-	#176	TOWEL GOURD	絲瓜	/*sz̩^{1} k^{w}a^{1}/

Table 6.3: Examples of the Common Yue Plain Stop Onsets

6.1.2 Aspirated Stops

Aspirated stops largely remain unchanged across the dialects. The chain shifts covered for the alveolar and post-alveolar series also applied to the aspirated onsets. However, in the case of NH and WC, the last part of the chain shift did not occur, leading to a merger of dental and alveolar aspirate stops.

Interestingly, in the whole Sìyì group CY /*t^{h}-/ evolved into /h-/. This initial change may have spread to other dental onsets with different laryngeal qualities and initiated a pull chain in the TS and KP dialects. Before syllabic /z̩/, cases of CY aspirated dental and avleolar affricates have merged intermediately as /*s-/ before CY /*s-/ further

participated in a chain shift in the TS and KP dialects and became /ɬ-/.

The aspirated velar onset is very sparsely attested in actual vocabulary. We have no clear examples of it in our reconstructed lexicon. This is likely because /*k^h-/ merged with /*x-/ in an early stage of the language, then was reintroduced as a phoneme upon contact with northern Chinese *koiné*. Where Common Yue has instances of the onset /*x-/ many non-Yuè dialects have reflexes of either /k^h-/ or /x-/. Coblin postulates a similar merger in Early Southern Highlands Chinese.

We have only one example of the aspirated labiovelar onset in etymon #084 CHOPSTICKS 筷 /*$k^{wh}aj^5$/. It is possible that the various reflexes of this word do not actually go back to a common origin. Because we do not have other examples of this onset, it is difficult to tell if the correspondence in the case of #084 CHOPSTICKS is stable.

Onset	Example			
*p^h-	#228	RIBS (2)		/*p^hian^1/
*t^h-	#004	LOOK	睇	/*t^haj^3/
*ts^h-	#117	SEVEN	七	/*ts^hit^7/
*$tʃ^h$-	#128	CHINA FIR	杉樹	/*$tʃ^ham^5$ $ʒy^6$/
*k^h-	[no clear examples]			
*k^{wh}-	#084	CHOPSTICKS	筷	/*$k^{wh}aj^5$/

Table 6.4: Examples of the Common Yue Aspirated Stop Onsets

6.2 Voiced Stops

None of the dialects in the current study preserve CY voiced obstruents in their original form. However, there are reports that some Yuè dialects in Guǎngxī preserve voiced

obstruents as voiced murmured obstruents, similar to many modern Wú dialects (Tsuji, 1980). There are two main trends in the development of voiced stops in Yuè dialects. First, in the Guǎngfǔ and Sìyì groups, CY voiced stops developed into aspirated stops in reflexes of tones 2 and 4 and plain stops in tones 6. Tone 8 is overwhelmingly plain as well, but some reflexes are aspirated, like in etymon #092 THIEF /*[dz]ɐk^{8}/. Second, in the Fēnglián and Yōngxún groups, voiced stops develop into plain stops unconditionally. Some of the Yōngxún dialects situated near urban centers, like the BL and QZ dialects follow mostly the Guǎngfǔ development of voiced stops. This suggests dialect mixture, likely caused by recent waves of Guǎngfǔ immigration to the area. The WC dialect follows neither of these trends. All CY voiced stops have evolved into aspirated stops.

Like with the aspirated affricates, CY /*dz-/ and /*dʒ-/ merged as /*s-/ before syllabic /*ʐ̩/ in reflexes of tone classes 2 and 4 in the Sìyì group. First, devoicing occurred, giving way to a wider pool of aspirated affricates. These affricates then merged in an intermediate stage with the fricative /*s-/ before syllabic /*ʐ̩/ before further diversifying in the chain shifts discussed above.

The NH dialect demonstrates both Guǎngfǔ-Sìyì and Fēnglián-Yōngxún trends due to accommodation of the Guǎngfǔ regional standard. The unaspirated forms suggest that the NH dialect and other Guǎngfǔ group dialects that border the northern inland region may once have shared closer affiliation with our Fēnglián group. We consider this to be a case of Malkiel's reapportioning of dialect zones (Malkiel, 1964, p.178). The northern Guǎngfǔ dialects' membership is due to a more recent shift in the dialect center of these

sites as the provincial seat in the Pearl River Delta became an important cultural hub. Yue-Hashimoto (1988; 1991) noticed this conservative trend of voiced stops evolving into plain stops in tones 2 and 4 rather than aspirated stops in the development of initials in our northern Guǎngfǔ dialects and placed them in her Northern Delta group on account of this. This classification is perhaps valid as a diachronic taxonomy, but it is clear from our measurements of linguistic distance that these dialects descriptively are more similar to the dialects around the Pearl River Delta region than they are the dialects of the northern inland.

Recent investigators have regarded character readings involving plain stops in tones 2 and 4 as "colloquial readings" and the aspirated stops in the same environments as "literary readings" owing to different lexical layers (Péng, 2005; Yue-Hashimoto, 2006). Péng (2005, pp.183–184) observes that many of the dialects of the area surrounding NH and the Zhàoqìng-Sānyì 肇慶三邑 region exhibit these layered readings, and draws comparisons to the layered "colloquial" plain and "literary" aspirated stops in lower *yáng* register tones in Shānxī and Shāndōng provinces. While he acknowledges the considerable distance between these locales, Péng implies in his conclusion that the earlier layer of plain stops and the later layer of aspirated stops come from common waves of linguistic regional standards spreading to these regions, stating "it cannot help but make one ponder deeply." Adopting a broader typological perspective, devoicing of voiced obstruents following plain and aspirated development paths in different tone classes may not be so ponderous after all. This series of developments has occurred in parallel across

hundreds of languages in the greater East Asia and Southeast Asia area. No doubt many disparate locales have exhibited common tone class and laryngeal feature patterning despite bearing no close relation to each other. The dialects to which Péng compares the Sānyì-Zhàoqìng subgroup only exhibited layering of plain and aspirated reflexes of historically voiced stops in tone class 2, not also tone class 4. Layering in tone 2 of these dialects was likely a parallel development instead of a common innovation.

We believe that this layering, as well as the layering of the consonant chain shift in the NH dialect described above, are the result of relatively recent dialect mixture rather than the result of the introduction of a northern literary *koiné*. [1] Layered reflexes occur not only in character readings, but also in a large amount of basic vocabulary items. The additional correspondences manufactured by this instance of dialect convergence are not represented in any of the other dialects of CY and are not projected back as separate distinctions in the common system.

Curiously, in the NH dialect we see instances of CY /*d-/ developing in three ways, /ʈ-/, /t^h-/ and /t-/, in tone class 2. While /t^h-/ is expected due to dialect leveling, a dichotomy of both /ʈ-/ and /t-/ is not. If the NH consonant chain shift occurred after CY voiced obstruent devoicing, we would expect only /ʈ-/. If it occurred before and CY /*d/ did not participate in the chain shift the same way CY /*t-/ did, then we would expect /*t-/. Péng (2005, p.199) explains the seemingly incomplete shift of ts >t >ʈ as

[1] "Relatively recent" likely still antedates the late 19th century. Dyer Ball's account of the Shùndé dialect of this region exhibits uniform aspiration of CY voiced onsets in reflexes of tones 2 and 4, i.e., the typical Guǎngfǔ development (Ball, 1900a,b).

a “compensatory” sound change.[2] /t-/ only evolved into /ʈ-/ in cases where /ts-/ pushed it so as not to create new homophone groups. In syllables that started in original onset /t-/, but had no minimal pair starting in onset /ts-/, /t-/ was retained because there was nothing to push it. Péng does not take a hard stance on whether this shift occurred after devoicing or before. In the paragraph, he makes reference of the shift occurring in the entire QYS dental stop series, *du ānzǔ* 端組, but goes on to say that the QYS aspirated dental stop *th-*(*tòumǔ* 透母) did not shift to a retroflex position, and only gives examples of the shift and discusses discrete development in terms of the QYS plain dental stop *t-*, (*duānmǔ* 端母).

We may reject Péng’s explanation on two accounts. First, his “compensatory” sound change proposal is antithetical to the modern understanding of regular sound change. Regular sound change is exceptionless. It may be conditioned by immediate phonetic environments, but is blind to the overall phonological inventory of a given language. It is not conditioned on a word-by-word basis to maximize overall syllable contrasts. Such a proposal does not conform with modern theories of lexical diffusion either. Second, there are an overwhelming number of character reading counterexamples to Péng’s proposed “compensation” that have indeed resulted in merger. We give a few examples in Table 6.6. The first of these examples cannot be explained as Guǎngfǔ influence because of the lack of aspiration in tone 2. So far, we have been unable to find any conditioning environment that prevents retroflexion and leave the matter for further study.

---

[2]語音演變中的 “補償” 現象

Voiced velar stops follow the same conditioned developments their plain counterparts did.

CY	b	b	d	d	dz	dz	dʒ	dʒ	g	g	gʷ	gʷ
Onset Cond.	2,4	6,8	2,4	6,8	2,4	6,8	2,4	6,8	2,4	6,8	2,4	6,8
YS	p	p	t	t	ts	ts	ts	ts	k	k	kʷ	kʷ
LX	p	p	t	t	ts	ts	tʃ	tʃ	k	k	kʷ	kʷ
HJ	p	p	t	t	ts	ts	ts	ts	k/ts	k/ts	kʷ	kʷ
FK	p	p	t	t	ts	ts	ts	ts	k/ts	k/ts	kʷ	kʷ
GZ	pʰ	p	tʰ	t	tsʰ	ts	tsʰ	ts	kʰ	k	kʷʰ	kʷ
NH	p/pʰ	p	ʈ/tʰ/(t)	ʈ/t	t/tsʰ	t/ts	ts/tsʰ	ts	k/kʰ	k	kʷ/kʷʰ	kʷ
DG	pʰ	p	tʰ	t	tsʰ	ts	tsʰ	ts	kʰ	k	kʷʰ	kʷ
FG	pʰ	p	tʰ	t	tsʰ	ts	tsʰ	ts	kʰ	k	kʷʰ	kʷ
YF	pʰ	p	tʰ	t	tsʰ	ts	tsʰ	ts	kʰ	k	kʷʰ	kʷ
XH	pʰ	p	h	t	tsʰ/s	ts	tsʰ/s	ts	kʰ	k	kʷʰ	kʷ
TS	pʰ	p	h	t	tʰ/ɬ	t	tsʰ/ɬ	ts	kʰ	k	kʰ	k
EP	pʰ	p	h	t	tsʰ/s	ts	tsʰ/s	ts	kʰ	k	kʰ	k
KP	h	p	h	t	tʰ/ɬ	t	tsʰ/ɬ	ts	kʰ	k	kʰ	k
BL	pʰ	p	tʰ	t	tʃʰ	tʃ	tʃʰ	tʃ	kʰ	k	kʷ	kʷ
QZ	pʰ	p	tʰ	t	tʃʰ	tʃ	tʃʰ	tʃ	kʰ	k	kʷ	kʷ
NN_p	p	p	t	t	tʃ	tʃ	ts	tʃ	k	k	kʷ	kʷ
BS_p	p	p	t	t	tʃ	tʃ	ts	tʃ	k	k	kʷ	kʷ
BY	p	p	t	t	tʃ	tʃ	tʃ	tʃ	k/tʃ	k/tʃ	kʷ	kʷ
MS	p	p	t	t	ɬ	ɬ	tʃ	tʃ	k/tʃ	k/tʃ	kʷ	kʷ
LX	pʰ	p	tʰ	t	tʃʰ	tʃ	tʃʰ	tʃ	kʰ	k	kʷʰ	kʷ
WC	pʰ	pʰ	tʰ	tʰ	tʰ	tʰ	tʃʰ	tʃʰ	kʰ	kʰ	kʷʰ	kʷʰ

Table 6.5: Common Yue Voiced Stop Onsets

Character	藤	層	笛	席	墊	箭
CY	*dɐŋ^{2}	*dzɐŋ^{2}	*diak^{8}	*dziak^{8}	*tian^{5}	*tsian^{5}
NH	tɐn^{2}	tɐn^{2}	tɛk^{8}	tɛk^{8}	tin^{5}	tin^{5}

Table 6.6: NH t/ts Mergers

Onset	Example			
*b-	#223	NOSE		/*bi^{6}/
*d-	#214	HEAD	頭	/*dɐw^{2}/
*dz	#221	NAVEL	臍	/*dzi^{2}/
*dʒ-	#098	DUST	塵	/*dʒin^{2}/
*g-	#181	3sg	佢	/*gy^{4a}/
*g^{w}-	[no clear examples]			

Table 6.7: Examples of the Common Yue Voiced Stop Onsets

6.3 Fricatives

6.3.1 Plain Fricatives

CY Onset Cond.	f	s	ʃ /_ʐ̩	ʃ	x /_u{j,n,t}	x /_u#	x^w /_C	x	x^w
YS	f	s	s	s	f	f	f	h	f
LX	f	s	ʃ	ʃ	f	f	f	h	f
HJ	f	θ	θ	θ	f	f	f	h	f
FK	f	s	s	s	f	f	f	h	f
GZ	f	s	s	s	f	f	f	h	f
NH	p^h	s	s	s	w/h	f	f	h	w
DG	f	s	s	s	f	f	f	h	f
FG	f	θ/s	s	s	f	f	f	h	f
YF	f	s	s	s	f	f	f	h	f
XH	f	s	s	s	f	f	f	h	f
TS	f	ɬ	ɬ	s	f	f	f	h	f
EP	f	s	s	s	f	f	f	h	f
KP	f	ɬ	ɬ	s	f	f	f	h	f
BL	f	ɬ	ʃ	ʃ	w	h	h/w	h	w
QZ	f	ɬ	ʃ	ʃ	f	f	f	h	f
NN_p	f	ɬ	ɬ	ɬ	w	h	h	h	w
BS_p	f	ɬ	ʃ	ʃ	w	h	h	h	w
BY	f	ɬ	ʃ	ʃ	w	h	h	h	w
MS	f	ɬ	ʃ	ʃ	w/f	h	h	h	w
YJ	f	ɬ	ʃ	ʃ	f	f	f	h	f
WC	f	ɬ	ʃ	ʃ	f	f	f	h	f

Table 6.8: Plain Fricative Onsets

The plain fricatives exhibit similar patterns of change as the other obstruents. CY /*s-/ has shifted to /*ɬ-/ in all of the Yōngxún group and in the MS, TS, KP, YJ, and WC dialects. In the FG dialect, it shifted to /θ-/. In this way, the distinction between CY /*s-/

and /*ʃ-/ was preserved in a considerable number of dialects.

/*s-/ and /*ʃ-/ merged as /ɬ-/ before syllabic /z̩/, in the TS and KP dialects and developed in to /ɬu/. The same is true of the FG dialect. Occasionally, /s-/ replaces regular /θ-/, likely due to surrounding Guǎngfǔ influence.

Labiodentals versus Labiovelars

The phonemes /*x^w-/ and /*f-/ of CY are not distinguished in the reconstructions of McCoy or Tsuji, although Tsuji did reconstruct a difference between /*f-/ and /*h-/ before his final /*-ow/ (CY /*-u/). Yue-Hashimoto acknowledges separate sources of modern labiodental fricatives, but does not draw the boundary between these sources in the same place that we do in CY. We may draw the distinction between /x^w/ and /*f/ in the common system on the basis of the NH dialect and the NN_p, BS_p, and BY dialects of the Yōngxún group.

CY /*f-/ is set up on the basis of correspondence with NH /p^h-/ and cases of /f-/ in the Yōngxún group. CY /*x-/ developed into /f-/ before /*u/ in dialects outside of the Yōngxún group in all environments but in syllables with the nucleus /*u/ that ended in velar codas. This is likely because /*u/ was phonetically lax before velar codas in CY just as it is in the majority of Yuè dialects today. Phonetically tense CY /*u/ conditioned rounding of /*x-/ that eventually evolved into /f-/, just as labiovelar /*x^w/ also overwhelmingly became /f-/. The NH dialect and the Yōngxún group dialects are the exception to this rule.

In the NH dialect and some of the dialects in the Zhàoqìng-Sānyì area as well as the BL, NN_p, BS_p, and BY dialects of the Yōngxún group, CY /*x-/ merged with /*x^w-/ in certain environments where vowel breaking occurred. The most uniform of these is before acute codes. In this environment, CY /*xuT/ developed into /x^wɐT/. Then, in this intermediate stage, /*x^w-/ developed into /w-/, leaving it distinct from CY /*f-/. The Yōngxún dialects also preserved /*x-/ as /h-/ before /*u/. This also applies to /*x^w-/ before /*-ɔ/ in open syllables, which raised to /-u/ in this environment in these dialects. However, in these cases, NH and the dialects that surround in followed the larger Guǎngfǔ trend and /*x-/ and /*x^w-/ developed into /f-/.

Curiously, CY /*f-/ corresponds to /p^h-/ in the NH dialect. This does not seem like a plausible sound change and suggests that this distinction in the common system is an artifact of northern *koiné* loans around the 12th century. Our choice to represent /*x^w-/ as separate from /*f-/ is not so much a statement on the actual separate articulation of these phonemes at a discrete time in the past for the entire geographic area so much as it is an abstraction to accommodate regular correspondence in NH, BL, NN_p, BS_p, and BY that necessitates separate phonemes in the common system. Likely, neither the /*f/ nor /*v/ of the CY system were present in the phonological system of the stage of language that preceded CY before the introduction of *koiné* loans. Before the introduction of *koiné* loans, /*x-/ before /u/ probably shifted to a labiodental pronunciation in the urban dialects around the Pearl River Delta, making it suitable to render northern /f-/. However, in the outlying dialects that change was either not in progress or had only partially completed,

as in the case of the NH dialect. Instead, these dialects either substituted an existing phoneme, like /p^h-/ in NH, to compensate, or innovated a completely new phoneme like /f-/ in the Yōngxún group.

/*v-/ was completely foreign to the pre-CY phonological system, created entirely by analogy to account for the voiced counterpart of the new phonemic /*f-/. This resembles the origin of phonemic /v/ in English, originally only an intervocalic allophone of /f/ that acquired phonemic status when it was used word initially to render French loans. NH dialect rendered foreign /v-/ as /p^h-/ in all the same environments as described above. It used /p^h-/ to render the labiodental articulation, accounting for the voicing difference solely by tone. Presumably, of the redundantly marked features of voice and tone register, tone register held greater contrast for the speakers of NH dialect at the time and possibly for CY as a whole. Before CY /*ɔ/, NH dialect /w/ had already become [v] phonetically. NH dialect used this phonetic expression instead of /p^h/ to render foreign /v/ before /ɔ/. Because CY /*w-/ is clearly distinguished from /*v-/ everywhere else, we need not set up two initials in the common system to account for the instances in which NH reflects /*v > f/ instead of /p^h/ for CY /*v/. They are in complementary distribution.

Ultimately, the distribution of CY /*f-/ and /*v-/ is restricted to *koiné* loans. This observation is crucial for the lexical seriation of Yuè dialects.

CY Onset Cond.	v	ʒ	ɣ /_ia	ɣ /_iɔ	ɣ
YS	f	ts/s	h	j	h
LX	f	tʃ/ʃ	h	j	h
HJ	f	ts	h	w	h
FK	f	ts	∅	∅	∅
GZ	f	s	j	j	h
NH	p^h/f	ts/s	h	j	h
DG	f	s	h	h	h
FG	f	s	h	∅	h
YF	f	s	j	j	h
XH	f	s	z	z	h
TS	f	s	z/h	z	h
EP	f	s	z	z	h
KP	f	s	z	z	h
BL	f	ʃ	h	j	h
QZ	f	ʃ	j	h	h
NN_p	f	ɬ	h	w	h/j
BS_P	f	ʃ	h	h	h/j
BY	f	ʃ	j	j	h/j
MS	f	tʃ	h	h	h/j
YJ	f	tʃ	j	j	h
WC	f	ʃ	j/h	w	h

Table 6.9: CY Voiced Fricatives

6.3.2 Voiced Fricatives

CY /*ʒ/ contrasts with /*dʒ/ firmly in the lexicon despite a lack of a similar contrast between /*dz-/ and /*z-/. In the Fēnglián group and the MS dialect, CY /*ʒ-/ has mostly merged with /*dʒ-/. There are a few sporadic cases where YS and LX have it separate as a fricative, but these might be due to more recent contact rather than a retention. In the

etymon #121 10 十 /*ʒ'ip^{8}/ there are cases where all dialects have the correspondences for /*ʃ/ but in low register tones. We might be able to set up a separate phoneme /*ʒ'-/ for this correspondence, but it is not clear if there is something special about the conditioning environment /*-ip/ that might put it in complementary distribution with /*ʒ-/. A similar set of correspondences seems to appear in one of the words we see in the unreconstructable etymon #177 WHEAT AWN, but that form of the word is not widespread enough to set up correspondences, so we do not posit full correspondences for /*z-/

CY /*ɣ-/ has palatalized and merged with /*j-/ before /*i/ before complex vowel nuclei in many dialects. The only example of this if our lexicon outside of character readings is etymon #129 AMARANTH (1) 莧菜 /*ɣian^{6} ts^{h}ɔj^{5}/. /*ɣ-/ does not contrast with /*j-/ before simple /*i/. We reconstruct a /*j-/ onset in those cases since the correspondences are more similar to general development of /*j-/. Similarly, /*ɣ-/ and /*w-/ do not contrast before the nuclei /*u/ and /*uɔ/. We reconstruct /*w-/ before these nuclei since the common system phoneme in these environments does not develop in parallel with /*x-/. Exceptions are the finals /*-uŋ/ and /*-uk/ where the correspondences line up with other voiced velar fricatives. Additionally, CY /*ɣ-/ has also become /j-/ before the modern final /-a/ in the NN_p, BS_p, BY, and MS dialects.

Onset	Examples			
*f-	#111	WIND	風	/*fuŋ^{1}/
*s-	#114	THREE	三	/*sam^{1}/
*ʃ	#192	BODY	身	/*ʃin^{1}/
*x-	#240	VULVA	閪	/*xɐj^{1}/
*x^{w}-	#099	FIRE	火	/*x^{w}ɔ^{3}/
*v-	#091	STEAMED RICE	飯	/*van^{6}/
*ʒ-	#219	LIPS	脣	/*ʒin^{2}/
*ɣ-	#080	ALLEY	巷	/*ɣɔŋ^{6}/
	#129	AMARANTH (2)	莧菜	/*ɣian^{6} ts^{h}ɔj^{5}/

Table 6.10: Examples of the Common Yue Fricative Onsets

6.4 Nasals

In Table 6.11 we do not give full correspondences for both plain and pre-glottalized sonorants because they only differ in tone. The reason we reconstruct pre-glottalized sonorants instead of voiceless sonorants like Norman and Coblin did is because 1. it seems typologically unlikely that /x^{w}/ would contrast with /w̥/(Kehrein and Golston, 2004). 2. there is no modern tendency of lenition and development into fricatives of these initials found in any Yuè dialect and 3. "voicelessness" of sonorants is actually a case of the laryngeal feature [spread glottis] which did not condition the tone split elsewhere. Aspirated obstruents of Common Yue are also defined by the feature [spread glottis], but it is not this feature that conditioned the tone split. It is the feature [voice] that conditioned the tone split. Pre-glottalized sonorants, marked by the feature [constricted glottis], would presumably have had a delayed voice onset time phonetically similar to the plain and aspirated obstruents and voiceless fricatives. The voiced obstruent and plain sonorants

would have carried over voicing from previous utterances, giving them a negative voice onset time in connected speech that in turn lowered tone register.

This judgment is not without its own problems, though. Languages that have pre-glottalized sonorants are usually complemented with a set of glottalized obstruent onsets like ejectives or implosives (Maddieson, 1984). This would make CY typologically unusual, but not entirely implausible. To avoid this unusualness, one could eliminate the onset distinction entirely and recast the distinction as a difference in tone class in the common system. It may have been the case that even during the early stages of diversification of the Yuè dialects, the distinction of tone register had more phonologically distinctive power than laryngeal contrasts involving voicing in onsets. However, this would complicate the development of aspiration in the obstruent onsets. For now, we will maintain that the original distinction was pre-glottalization.

Mirroring the obstruents, the nasal onsets of CY remain mostly stable except for the palatal onset. The palatal nasal merged /*ɲ-/ with /*j-/ in most of the Guǎngfǔ group and is reflected as either a velar nasal or palatal nasal elsewhere. The hardening to a velar nasal is typical of the Sìyì group. The DG dialect is unique in that the one pre-glottalized reflex of this onset that we have, in etymon #202 EARS 鼻 /*ˀɲi^{3}/, preserves it as a nasal, while plain reflexes are typically reflected as /z-/. It is unclear if pre-glottalization conditioned this split or it is the result of dialect mixture.

CY /*ŋ-/ palatalized and merged with /*ɲ-/ in most dialects before close front vowels. The exception is the LX dialect where they remain distinct like in the etyma #100 MOON

月 /*ŋiɔt^{8}/ and #032 FISH 魚 /*ŋy^{2}/.

The CY onset /*ŋ^{w}-/ is very marginal. It is recoverable on the bases of the NN_p, BS_p where it lenited to /*w-/ before /*ɔ/ and on the BY and MS dialects where it is preserved in its original state. Glimpses of this onset are also preserved in dialects when it occurs before the nucleus /*ɔ/. In this environment, /ɔ/ follows the trajectory of the development of the nucleus /uɔ/ in some dialects and is kept separate.

CY Onset Cond	m	n	ɲ	ŋ [1]	ŋ	ŋ^{w}	l	j [2]	j [3]	j	w [4]	w
YS	m	n	j	j	ŋ	ŋ	l	j	j	j	w	w
LX	m	n	ɲ	ŋ	ŋ	ŋ	l	j	j	j	w	w
HJ	m	n	ɲ	ɲ	ŋ	ŋ	l	j	j	j	w	w
FK	m	n	ɲ	ɲ	ŋ	ŋ	l	j	j	j	w	w
GZ	m	n	j	j	ŋ	ŋ	l	j	j	j	w	w
NH	m	n	j	j	ŋ	ŋ	l	j	h	h	w	f/w
DG	m	n	z/ŋ	z/ŋ	ŋ	ŋ	ŋ	z	z	z	v	v/f
FG	m	n	j	j	ŋ	ŋ	l	j	∅	j	w	w
YF	m	n	j	j	ŋ	ŋ	l	j	j	j	w	w
XH	m	n	ŋ	ŋ	ŋ	ŋ	l	z	z	z	v	v
TS	m	n	ŋ	ŋ	ŋ	ŋ	l	z	z	z	v	v
EP	m	n	ŋ	ŋ	ŋ	ŋ	l	z	z	z	v	v
KP	m	n	ŋ	ŋ	ŋ	ŋ	l	z	z	z	v	v
BL	m	n	ɲ	ɲ	ŋ	ŋ	l	j	j	j	w	w
QZ	m	n	ɲ	ɲ	ŋ	ŋ	l	j	j	j	w	w
NN_p	m	n	ɲ	ɲ/w	ŋ	w	l	j	w	h	w	h
BS_p	m	n	ɲ	ɲ/w	ŋ	w	l	j	w	h/j	w	h
BY	m	n	ɲ	ɲ	ŋ	ŋ^{w}	l	j	j	j/h	w	h/w
MS	m	n	ɲ	ɲ	ŋ	ŋ(u)	l	j	j	h	w	h
YJ	m	n	j	j	ŋ	ŋ	l	j	j	j	w	w
WC	m	n	ɲ	ɲ/w	ŋ	ŋ	l	j	w	j	w	w

[1] /_{i,y} [2] /_i{m,p,n,t,w} [3] /_ɔ{T} [4] /_u{n,t,j}

Table 6.11: Common Yue Sonorants

Onset	Examples			
*m-	#025	CUTTLEFISH	墨魚	/*mɐk^{8} ŋy^{2}/
*n-	#101	MUD	泥	/*nɐj^{2}/
*ɲ	#113	TWO	二	/*ɲi^{6}/
*ŋ-	#156	LOTUS	藕	/*ŋɐw^{4b}/
*ŋ^{w}-	#072	in-law/maternal	外	/*ŋ^{w}ɔj^{6}/
*ʔm-	#060	TAIL	尾	/*ʔmi^{3}/
*ʔn-	#195	BREASTS		/*ʔnɛn^{1}/
*ʔɲ-	#202	EARS	耳	/*ʔɲi^{3}/
*ʔŋ-	#203	EYE	眼	/*ʔŋan^{3}/
*ʔŋ^{w}-	#093	TILE	瓦	/*ʔŋ^{w}a^{3}/

Table 6.12: Examples of the Common Yue Nasal Onsets

6.5 Approximants

CY /*l-/ is the most stable of the approximants. It has merged with /ŋ/ in the DG dialect. In many dialects outside of this study, /n-/ has merged with /l-/.

The situation with the semivowels is more complicated. In the NH, NN_p, BS_p, and MS dialects /*j-/ merged with /*ɣ-/ unconditionally, but later changes due to vowel breaking shifted the onset back to /j-/. Following the trend of the general Yuè vowel shift, the nucleus /*i/ broke to to /*iɐ/ in an intermediate stage before codas involving stricture in the mouth. Before this, new /*iɐ/, /*ɣ-/ became /*j/ in the NH, NN_p, BS_p, and MS dialects leaving just /jɐ/. In the BS_p dialect, a further break of nucleus /*i/ to /*iə/ occurred before velar codas, producing a similar shift. In the BY dialect, /*j-/ also shifted to /*ɣ-/, but only before the nucleus /*y/.

The development of CY /*w-/ followed a very similar trajectory in the same dialects, NH, NN_p, BS_p, and MS, but also seems to have partially occurred in DG and BY. CY /*w-/ shifted to /*ɣ-/, but in cases where CY /*u/ broke into /uɐ/, it shifted back to /w-/.

In the NN_p, BS_p, and WC dialects, CY /*jɔ/ before acute codas changed to /jy/, as it did in many other places, but /y/ preceded to break into /wi/. In these cases, /w/- is now the onset instead of /j-/.

Semivowels have fricativized in the DG dialect and the dialects of the Sìyì group.

Onset	Examples			
*l-	#064	WOLF	狼	/*lɔŋ^{2}/
*j-	#089	SALT	鹽	/*jam^{2}/
*w	#097	CLOUD	雲	/*wun^{2}/
*ʔl-	#161	OLIVE	欖	/*ʔlam^{3}/
*ʔj-	#033	FLY	蠅	/*[ʔj]iŋ^{1}/
*ʔw-	[no clear examples]			

Table 6.13: Examples of the Common Yue Approximant Onsets

6.6 The Zero Onset and the Phonemic Status of Laryngeal and Semivowel Onsets

The zero onset of CY remains zero in the dialects except before the close vowels, /i,y,u/, in most dialects. For the front vowels, it was reinforced as /j-/, and for the back rounded vowel, it was reinforced as /w-/. Rather than project pre-glottalized approximants back in these cases, we treat this as a conditioned change of the zero onset.

There is no contrast of semivowels occurring before similar vocalic segments in complex nuclei. That is to say, /jia/ does not contrast with /ja/, nor /wɔ/ with /wuɔ/, etc. Because of this interesting phonological gap, it might seem tempting to recast these onsets as one low register zero phoneme, /*ɦ/, and rewrite /ja, ji, jy, wu, wɔ/ as /ɦia, ɦi, ɦu, ɦuɔ/. This is not possible, though, because the onset /*w-/ occurs before/*a, *i, *ia/ and contrasts directly with /*j-/.

Similar to how the semivowels do not occur before similar vocalic segments in complex vowel nuclei, labiovelars generally do not occur before /uɔ/. We treat them as sequences /K^{w}ɔ/ instead of /Kuɔ/ or /K^{w}uɔ/. It is important to note, though, that often times /*ɔ/

after labiovelar onsets develops as CY /*uɔ/ would.

6.7 Major Differences with Other Reconstructions

Besides just the phonemes in the inventory, our system's correspondences differ from Yue-Hashimoto (2006) in a few key ways.

First, we treat the development of the labiodental fricatives as the result of rendering foreign phonemes /*f-/ and /*v-/. Yue-Hashimoto (2006) treats their development with both a primary source, /*β/ that could occur in either the high *yīn* and low *yáng* registers of tone, and a secondary source as a conditioned change of bilabial stops before the high back unrounded vowel (Yue-Hashimoto, 2006, p.99). The latter proposal is not supported by regular correspondence. Many dialects reflect layered "character readings" containing both bilabial plosives and labiodental fricatives, deriving from words of differing historical origin. Such correspondences are erratic. The stop readings and fricative readings do not occur consistently in the same dialects across "cognate" readings. In fact, many of the dialects have both "readings," reflecting two different words written using a common character. It is not possible to predict which dialects retained the bilabial plosive "reading" and which innovated the labiodental fricative "reading" consistently. It ultimately is a matter of which word is being said that decides the initial of the reading of the character used to write that word and whether or not such words have been displaced among the dialects. Take, for instance, the etymon #091 STEAMED RICE /*van^{6}/. The inland LX, FK, and MS dialects have a bilabial plosive reading, while a labiodental fricative is

reflected elsewhere. In other cases of labiodental onset words, these three dialects do not consistently have bilabial stops. It is just that they preserved an older word for steamed rice than the other dialects. It is also possible for other dialects outside these three to have retained bilabial stop "readings." In the variant etyma #67 and #68 DAUGHTER-IN-LAW 新婦 /*sin^{1} bu^{4a} ~fu^{4a}/ nearly all the dialects in the Fēnglián, Guǎngfǔ and Sìyì dialects have bilabial stop onsets readings of the character 婦, but in the Yōngxún group it is predominantly read with a labiodental fricative onset. The correspondences are not consistent, so we cannot propose a conditioned changed of the bilabial stops into labiodental fricatives. It is better to not suppose any sort of conditioned change of biblabial stops in Yuè dialects.

McCoy (1966) did not treat labiodental fricatives as the result of conditioned sound change, but he also did not distinguish them from CY's plain labiovelar fricative of conditioned change of velar fricatives. Tsuji (1980) also projected labiodental fricatives back to the proto-system, but only distinguished the two sets of fricatives before certain finals.

Like Yue-Hashimoto, Péng (2005, pp.186–191) regards modern labiodental fricatives as the result of conditioned sound change and the NH dialect's instances of /p^{h}/ as instances of retention of former bilabial stops. Like Yue-Hashimoto, he suggests that bilabial stops fricativized before medials or nuclei that contained /u/ using corresponding QYS terminology, but went through an intermediate stage where they were realized as /p^{h}-/. He then suggests more or less the same divide as we do for where labiodentals elsewhere are realized as /f-/ instead of /p^{h}-/ in terms of corresponding QYS rime categories, but

frames this as another conditioned change rather than an issue of matter of borrowing foreign sounds. His explanation is not grounded in discrete phonological contrasts in the ancestor of Yuè or even in discrete QYS reconstructions making it impossible to evaluate its plausibility. Listing rime table categories without discussing what actual phonological features underlie them and conditioned change renders them arbitrary. It frames otherwise readily observable patterns in an abstruse fashion without explaining why the patterns formed. Péng does not account for why other dialects occasionally retain bilabial stop readings with rigorous correspondence, either. His proposal for conditioned changes should be rejected on the same grounds as discussed above.

Yue-Hashimoto (2006, pp.128–130) only reconstructs semivowel approximants in the upper *yīn* register tones. We reconstruct the zero onset in these cases, treating the semivowels as conditioned reinforcement of vowel nuclei as discussed above. She treats the Common Yuè semivowel /*w-/ in lower *yáng* register tones as conditioned change of /*ɦ-/ before /*u/ in medial of nucleus position. This is similar to the way we handle the zero onset before {i,y,u}, but applied to the lower register in just the case of /*u/. We reject medial /u/ completely, so all cases of /*ɦu-/ in here system are /*w-/ in ours. Her /*ɦ-/ in other cases corresponds to our /*ɣ-/. She reconstructs /*ɣ-/ where we reconstruct /*j-/ in the cases of the /h-/ reflexes in her Sānyì-Zhàoqìng group, i.e. where the NH dialect has them in our correspondences. It is more economic to deal with these reflexes as innovations rather than retentions. Those cases of /*j-/ becoming /h/ are in complementary distribution with where we reconstruct /*j-/ elsewhere.

Chapter 7

CODAS

7.1 Interaction Between Codas and Nuclei

The development of CY codas and nuclei are best understood in terms of the minimally contrastive features of each that conditioned change. For the codas, the contrasts are similar to the onsets except laryngeal features are no longer distinctive. Only manner and place contrast. A broad distinction between codas that involve stricture at the front the oral cavity, /-w,-m,-p,-j,-n,-t/ and codas that do not, /∅,-ŋ,-k/, is also helpful.

We may define the vowel nuclei of the majority of modern Yuè dialects by just three privative features, [front], [round], and [open]. Complex vowel nuclei involve a shift from [-open] to [+open] vowels, while phasing of [front] and [round] features across the diphthong is not distinctive. These features have conditioned coda innovations in a few of the dialects in our study, but most of these developments occurred after a widespread vowel chain shift before front stricture codas. Because of this, when we discuss coda development, it is more expedient to talk about vowel features post-shift rather than from the perspective of the original Common Yue features.

In the closer tenser front stricture coda environment, CY complex vowel nuclei predominantly monophthongized as mid vowels with like vocalic features, [front] and [round].

	/[+front/] /[-round/]	/[+front] /[+round/]	/[-front/] /[-round/]	/[-front/] /[+round/]
/[-open/]	i	y	ɐ	u
/[open/]	ɛ	œ	a	ɔ
(complex)	ia	iɔ		uɔ

Table 7.1: Vowel Nuclei Contrasts in Yuè Dialects

Then, peripheral [-open] vowel nuclei, /*i,*y,*u/ fell along the nonperipheral track, pulling the new mid vowels upwards to occupy their old spots. This chain shift was only partial in the Sìyì group and some dialects of the Fēnglián group. Each demonstrate different stages of progression. For the majority of dialects, we generalize this trend below. Where slashes are given, the form before the slash occurs after grave onsets including those labial, velar, labiovelar, and zero in place. The form after the slash occurs after the remaining acute onsets. Commas separate expressions found in different dialects. /j/ in parentheses indicates the development of onset /j-/ in the case of zero onset. /w/ in parentheses indicates development of labiovelar onsets to replace previous velar or zero onsets.

- (ia >) e →i →(j)ɐ
- (iɔ >) ø →y →(j)œ,(j)ɐ
- (uɔ >) o →u →(w)ɐ/y,u

Coda Cond.	w /ɔ_	w	j	m	n	ŋ	p	t	k
YS	∅	w	j/∅	m/n	n	ŋ	p/t	t	k
LX	w	w	j/∅	m/n	n	ŋ	p/t	t	k
HJ	w	w	j	m/n	n	ŋ	p/t	t	k
FK	∅	w	j/∅	m	n	ŋ	p	t	k
GZ	w	w	j	m	n	ŋ	p	t	k
NH	∅	w	j/∅	m	n/ŋ	ŋ	p	t/k	k
DG	w	w	j	m/n/ŋ	n/ŋ	ŋ	t/k/∅	t/k/∅	k/∅
FG	w	w	j	m	n	ŋ	p	t	k
YF	w	w	j/∅	m	n/ŋ	ŋ	p	t/k	k
XH	∅	w	j	m	n	ŋ	p	t	k
TS	w	w	j	m	n	ŋ/n	p	t	k/t
EP	w	w/j	j	m	n	ŋ	p	t	k
KP	∅	w	j	m	n/ŋ	ŋ	p	t/k	k
BL	w	w	j	m	n	ŋ	p	t	k
QZ	w	w	j	m	n	ŋ	p	t	k
NN_p	w	w	j	m	n	ŋ	p	t	k
BS_p	w	w	j	m	n	ŋ	p	t	k/t
BY	w	w	j/∅	m	n	ŋ	p	t	k
MS	w	w	j	m	n	ŋ	p	t	k
YJ	w	w	j	m	n	ŋ	p	t	k
WC	w	w	j	m/n	n	ŋ	p/t	t	k

Table 7.2: Common Yue Codas

7.2 Correspondence and Development of the Codas

Table 7.2 demonstrates the basic correspondence of the codas. Generally, modern Yuè dialects tend to exhibit conservation of historical codas. Coda innovations are mainly conditioned by the features [open] and [front]. Of all the dialects, the DG is the most innovative in terms of codas. We will set it aside for the time being and devote a separate

subsection to its development.

7.3 Approximant Codas

A few dialects in the Fēnglián, Guǎngfǔ, and Sìyì groups lost the /*-w/ coda after the open round vowel /ɔ/. Additionally, /*-w/ evolved into /*-j/ after /*ɐ/ in the EP dialect.

The YF dialects lost the coda /*-j/ after the nuclei /*ɔ/ and /*uɔ/. It also lost /*-j/ after /*u/ in syllables that begin in acute onsets. In this same environment /*uj/ coalesced into /y/ in the LX and YS dialects and into /œ/ in the FK and BY dialects. Additionally /*ɔj/ also coalesced into /œ/ in the BY dialects after grave onsets. The NH dialect lost the /*-j/ coda after all open nuclei and after /*u/.

7.4 Bilabial Codas

The YS, LX, and HJ dialects have merged bilabial codas with dental codas after post-shift front vowels, i.e. present day /i/ and /ɛ ~e/. The WC dialect has merged bilabial codas with dental codas after the present day rounded vowel /ɔ/. This has two historical origins, /*ɔ/ and /*ɐ/. In the case of the latter, Lǐ (2014) contradicts himself. The only etymon with /ɐ/ and a bilabial coda, #035 FROG 蛤 /*kɐp^{7}/ , is listed as /kɔk^{9}/ in the vocabulary section of the book, but as /kɔt^{9}/ in the character reading section of the book. All other instances of /*ɔ/ plus a bilabial coda evolve to /*ɔ/ plus a dental stop. While it is not impossible that CY /*ɐ/ conditioned the bilabial coda to become velar, it seems a bit unlikely. Therefore, with some hesitation, we tentatively consider the /kɔk^{9}/ transcription to be in error.

7.5 Dental Codas

The YF dialect merged dental and velar codas after the modern open vowel /a/. The NH dialect merged dental and velar codas after all modern open vowels, {ɛ,a,ɔ}. The KP dialect merged dental codas with velar codas after Common Yue /*ɛ/ and cases of Common Yue /*ia/ after the palatal onset.

7.6 Velar Codas

The TS dialect has merged velar codas with dental codas after modern /e/. The BS_p dialect has also merged velar codas with dental codas, but the condition is after CY /*i/ and /*ia/, which merged as /*i/ in an intermediate stage still preserved in the neighboring dialects of the Yōngxún group.

Coda	Examples			
*-w	#061	TIGER	老虎	/*lɔw^{4a} xu^{3}/
	#120	NINE	九	/kiw^{3}/
*-j	#135	BARNYARD GRASS	稗	/*bɐj^{6}/
	#220	MOUTH	嘴	/*tsuj^{3}/
*-m	#058	SILKWORM	蠶	/*dzam^{2}/
	#160	NONGLUTINOUS RICE	黏米	/*tʃiam^{1} ʔmɐj^{3}/
*-n	#184	DRY	乾	/*kɔn^{1}/
	#192	BODY	身	/*ʃin^{1}/
*-ŋ	#222	NECK	頸	/*kiaŋ^{3}/
	#010	STROLL	蕩	/*dɔŋ^{6}/
*-p	#017	BUTTERFLY	蝴蝶	/*wu^{2} diap^{8}/
	#028	DUCK	鴨	/*ap^{7}/
*-t	#042	Llouse	蝨	/*ʃit^{7}/
	#100	MOON	月	/*ŋiɔt^{8}/
*-k	#106	STONE	石頭	/*ʒiak^{8} dɐw^{2}/
	#109	FALL (RAIN)	落	/*lɔk^{8}/

Table 7.3: Examples of the Common Yue Codas

7.7 The Dōngguǎn Dialect

The DG dialect's codas diverge the most from CY. This is because codas in the DG dialect underwent two separate series of changes. One involved all nasal and stop codas. The other involved only stop codas.

The first series of changes started after the development of tone 9 was conditioned by CY open vowel nuclei and the general Yuè vowel chain shift had completed. First, dental stops evolved into velar stops after open vowel nuclei. At this point in time, only CY /*a/ remained open. Before dental codas, CY /*ɔ/ had merged with CY /*uɔ/ and had become /u/. After this change, all cases of /*-aŋ/ shifted forwards and became /-ɛŋ/.

Next, bilabial codas also evolved into velar codas after [open] vowel nuclei. CY /*a/ and /*ɔ/ had already merged into just /*a/ before bilabial codas, leaving just one change of /*-am/ to /*-aŋ/. Finally, the remaining bilabial codas merged with dental codas after front vowel nuclei.

In the second series of changes, dental and bilabial codas became velar codas after the vowel /ɐ/. Intermediate /*-ɐt/ became /-ɐk/ and /*-ɐp/ became /-ɔk/. Then, /*iɔ/ monophthongized to /*ø/ before velar codas. /*ia/ before velar codas backed and rounded, taking its place. All cases of tone 9 in syllables ending in velar codas after non-peripheral open vowel nuclei, /ɛ/, /ɔ/, and /iɔ/ lost their final velar coda and developed a new tone traditionally described as *biànrù* 變入. Afterwards, remaining velars codas occurring with tone 9 were lost and tone 9 in these cases merged with tone 5. All remaining cases of old tone 9 merged with tone 8. We mark the *biànrù* tone as 9 in our broad transcription instead of a tenth tone for economy even though it does not correspond consistently with tone 9 in other Yuè dialects. Finally, /iɔ/ merged with /ø/. All changes are summarized below.

1. Tone 7 > Tone 9 / V[open]
2. General Yuè Vowel Chain Shift (see above)
3. ɔ > a / _m,p
4. C[dental] > C[velar] / V[open]_
5. a > ɛ /_{ŋ,k}
6. C[labial] > C[velar] / V[open]_

7. C[labial] > C[dental] / V[front]_

8. {t,p} > k / ɐ_

9. iɔ > ø / _{ŋ,k}

10. ia > iɔ / _{ŋ,k}

11. Tone 9 > Tone 10 / {ɛ,ɔ,iɔ}k

12. k > ∅ / _# (Tone 10)

13. Tone 9 > Tone 5 / V[open]k

14. k > ∅ / _# (Tone 5)

15. Tone 9 > Tone 8

16. iɔ > ø

17. (rewrite tone 10 as tone 9 since tone 9 was eliminated)

Chapter 8

NUCLEI

Syllable codas exerted a great conditioning influence upon the development of vocalic nuclei among the Yuè dialects. As mentioned earlier, the place of these codas phonetically altered the oral cavity, influencing the degree of openness and peripherality of vowel nuclei. These changes eventually became phonemicized in some cases and fed into larger patterns of chain shifts.

8.1 The General Yuè Vowel Shift

We follow Labov et al. (2008)'s principles of vowel chain shift mechanics. In chain shifts,

1. Peripheral vowels (i.e. tense vowels) move upward along a peripheral track. Back vowels can move to the front at the top of this track.

2. Non-peripheral vowels (i.e. lax or centralized vowels) move downward along a nonperipheral track.

When vowels reach the terminus of their tracks in vowel space, they follow two exit principles,

1. The Lower Exit Principle: Low (i.e. open) nonperipheral vowels become peripheral.

e.g. /ɐ/ > /a/

2. The Upper Exit Principle: One of two high (i.e. close) peripheral morae becomes

non-peripheral.

e.g. /iː ~ij/ > [ɪj] or [iə]

Common Yue non-open vowels /*i, *y, *u/ and open /*a/ were peripheral. The non-open vowel /*ɐ/ and the open vowel /*ɔ/ were nonperipheral. The peripheral vowels were phonetically tense in all environments except for before velar codas, where they were likely phonetically lax, [ɪ, ʏ, ʊ].

In many Yuè dialects, in these phonetically tense environments, the non-open peripheral vowels broke into falling diphthongs in open syllables. /i/ became [ei], /y/ became [øy], and /u/ became [ou]. In some cases, these diphthongs gained phonemic status. We generally transcribe them as /ij/ and /uw/ in this study, but treat [øy] as /œj/. Presumably, these diphthongs progressed phonetically through a phase in which the former vowel was phonetically lax in accordance with the Upper Exit Principle, e.g. /*i/ > [*ɪi] > [*ei].

In other phonetically tense environments, i.e., before non-velar codas involving stricture at the front of the oral cavity, Common Yue non-open peripheral vowels overwhelmingly broke into rising diphthongs in accordance with the Upper Exit Principle.

- /*i/ > [iɐ] > /ɐ/ ~ /jɐ/
- /*y/ > [yɐ] > /ɵ/ ~ /jɵ/
- /*u/ > [y] ~ [uɐ] > /y/ ~ /wɐ/ ~ /ɐ/

These have now been reduced to simple nonperipheral vowels [ɐ] and [ɵ] in most dialects after nonzero onsets. /*i/ often retains the former part of the transitional diph-

thong as a phonemic /j-/ onset in cases of previous zero and laryngeal onsets. /*u/ did not follow as neat of a path as its front counterparts. /*u/ fronted to /y/ after acute onsets in some dialects after original /*y/ exited the upper peripheral track. Elsewhere, it broke as the other close peripheral vowels did. It retained the close round element as /w/ after zero onsets and conditioned velar onsets to become labiovelar onsets. This element was lost in other environments.

Before codas involving stricture at the front of the oral cavity, Common Yue complex vowel nuclei /*ia, *iɔ, *uɔ/ overwhelmingly monophthongized to mid, peripheral /e, ø, o/. These vowels took the place of the previous Common Yue non-open peripheral vowels, moving upward along the peripheral track. This change did not occur in the same fashion before zero or velar codas because the complex vowel nuclei were largely not monophthongized in these contexts.

We dub this series of changes the General Yuè Vowel Shift because most Yuè dialects except those on the furthest periphery demonstrate these changes.

8.2 Vowel Nuclei Before Zero and Velar Codas

Before zero and velar codas, vowel nuclei demonstrate the least amount of innovation.

/*ʐ̩/ only occurs after sibilant onsets in open syllables. The Guǎngfǔ group and Sìyì group demonstrate its contrast with /*i/the best. In those dialects, it seems to have first evolved into mainly evolved into /y/ before further development. Else where it has merged with /*i/.

Nucleus Cond.	ʐ̩	i /_#	i	y	u /ŋ_#	u /_#	u	ɐ	a	ɔ	ia	iɔ	uɔ
YS	i	ɐj/i	ɐ	œj/y	m̩	ɐw/u	o	ɐ	a/e	ɔ	i/e	œ	ɔ
LX	ɐj	ɐj/i	a	œj/y	m̩	ɐw/u	ɔ	a	a/ia	ɔ/œ	ɛ	œ	ɔw/ɔ/œ
HJ	ɐj	ɐj/i	ɛ	œj/y	ŋ̍	ɔw/u	u	ɐ	a/ɛ	ɔ	ɛ	œ	ɔ
FK	ɐj	ej/i	ɐ	œj/y	ŋ̍	u	ɔ	ɐ	a/e	uə/œ	i/iə	i	uə/–
GZ	i	ij/i	i	œj/y	ŋ̍	uw/u	u	ɐ	a	ɔ	ɛ	œ	ɔ
NH	y	ɐj/i	i	œj/y	ŋ̍	uw/u	u	a	a	ɔ	ɛ	œ	œ/ɔ
DG	ɐj	ɐj	i	uj/y	m̩	ɔw	u	ɐ	ɛ/a	ɔ	ø	ø	ɔ/u/ɔ
FG	y	ɐj/i	i	œj/y	ŋ̍	u	u	ɐ	a	ɔ	i/e	ø	ɔ
YF	i	ɐj/i	i	œj//y	ŋ̍	uw/u	u	a	a	ɔ	ɛ	œ	ɔ
XH	i	ij/i	i	uj/i	m̩	æw/u	u	a	a	ɔ	æ	iɔ	ɔ
TS	u	i	e	uj/i	m̩	u	u	a	a	ɔ	ia	ia	ɔ
EP	u	i	i	uj/i	m̩	u	u	a	a	ua	ia	iɔ	ua
KP	i	ij	i	uj/i	m̩	u	u	a	a	u/ɔ	ia	ia	u/ɔ
BL	i	i	i	y	ŋ̍	u	u	ɐ	a/ɛ	ɔ	ɛ	œ	o/ɔ/o
QZ	i	i	i	i	ŋ̍	u	u	ɐ	a	ɔ	ɛ	ɛ	ɔ
NN_p	i	uj/i	i	uj	ɔ	ɔ/u	u	ɐ	a/ɛ	a	i	ɛ	u/u/a
BS_p	i	uj/i	(j)ə	uj	ɔ	ɔ/u	ɔ	ɐ	a/ɛ	a	i/(j)ə	ɛ	u/u/a
BY	əj	əj/i	i	uj	əw	əw/u	u	ɐ	a/ɛ	œ/a	i	ɛ	əw/œ
MS	i	i	i	i	ŋ̍	uw/u	u	ɐ	a/ɛ	ɔ/ua	a/ɛ	iə	ɔ/ua/ɔ
YJ	ij	ij/i	i	i	uŋ	ɔw/u	u	ɐ	a	ɔ	ɛ	iɛ	ɔ
WC	i	i	i	i	u	u	u	ɐ	a	ɔ	i/ɛ	ɛ	ɛ/–

Table 8.1: Common Yue Nuclei Before Zero and Velar Codas

In open syllables, /*i/ broke into a rising diphthong in the Guǎngfǔ, Fēnglián, and inland part of the Yōngxún groups after all acute onsets that were not palatal or post-alveolar. In the GZ, FS, and YS dialects, this breaking also occurred after grave onsets. Diphthongization is also present in the XH and KP dialects. In the XH dialect, /*i/ has broken after acutes that are not palatal or post-alveolar. In the KP dialect, breaking is unconditioned in open syllables.

/*i/ remained mostly static before velar codas. The YS, HJ, FK, and ZH dialects demonstrate movement down the nonperipheral track conditioned by phonetic laxness. In the Sìyì group, /*i/ has merged with /e/ in the TS dialect, likely also conditioned by laxness before velar codas. In the BY dialect, /*i/ broke into /iə/ then simplified to /ə/ after zero or onsets that were not zero or semivowels. The codas after this new /ə/ shifted to a dental articulation.

In open syllables /*y/ largely followed the same pattern of breaking that /*i/ did. The Sìyì dialects present a more complicated state of affairs. CY /*y/ merged with /i/ after /j-/ and post-alveolar onsets that were not aspirated or did not develop aspiration after onset devoicing. In all other environments, /*y/ broke into /-uj/.

/*u/ also broke in open syllables after acute onsets that were not palatal or post-alveolar in similar dialects as listed above. Additionally, this breaking also occurred after labial onsets in the GZ, NH, XH dialects and the inland dialects of the Yōngxún group. The BY dialect broke /*u/ everywhere except after the onset /*w-/ The breaking of /*u/ simplified to /ɔ/ in the NN_p and BS_p dialects, likely as a result of the loss of a coda /-w/ after /ɔ/.

Before velar codas, /*u/ remains mostly unchanged among the dialects. The exception is the BS_p dialect where it seems to have broken and then simplified to /ɔ/, just as it did in open syllables.

/*ɐ/ only occurs in closed syllables. Before velar codas, this vowel has progressed downward along the nonperipheral track and merged with /a/ according to the Lower

Exit Principle in several dialects, most notably in the entire Sìyì group.

/*a/ in open syllables has remained unchanged in all dialects. Before velar codas, this vowel has fronted in the Fēnglián group, most of the Yōngxún group, and in the DG dialect.

/*ɔ/ has largely remained unchanged among the dialects. In the two "Pínghùa" varieties it has fallen to /a/ and fronted to /œ/ in the BY dialect.

/*ia/ has monophthongized to /ɛ/ in most dialects outside of the Sìyì group. In the inland Yōngxún group dialects, it has further risen to /i/.

/*iɔ/ is only marginally represented in open syllables. Our sole example is etymon #144 EGGPLANT (2) 茄 /*giɔ^{2}/. Before velar codas, it is largely unchanged apart from monophthongization and unconditional unrounding in some dialects.

The correspondences for /*uɔ/ when it occurs in open syllables and velar coda ending syllables are complicated and may not be stable since we only base them on a few cognates. In Table 8.1, the form before the first slash occurs in open syllables. The form after the second slash occurs before velar codas. The last form, when different from the second, occurs after bilabial onsets and before velar codas. For the last of the environments, only the reflexes of the BL and MS dialect distiguish the set of correspondences for cases of simple /*ɔ/.

Nucleus	Examples			
*ʐ̩	#126	INSTANCE MW	次	/*ts^{h5}ʐ̩/
*i	#012	BAT	飛鼠	/*fi^{1} ʃy^{3}/
	#241	WING	翼	/*jik^{8}/
*y	#032	FISH	魚	/*ŋy^{2}/
*u	#037	HARE	兔	/*t^{h}u^{5}/
	#151	GREEN ONION	蔥	/*ts^{h}uŋ^{1}/
*ɐ	#131	APRICOT	杏	/*ɣɐŋ^{2}/
	#92	THIEF	賊	/*dzɐk^{8}/
*a	#235	TEETH	牙	/*ŋa^{2}/
	#122	HUNDRED	百	/*pak^{7}/
*ɔ	#169	RADISH	蘿蔔	/*lɔ^{2} bak^{8}/
	#231	SHOULDER	膊頭	/*pɔk^{7} dɐw^{2}/
*ia	#059	SNAKE	蛇	/*ʒia^{2}/
	#106	STONE	石頭	/*ʒiak^{8} dɐw^{2}/
*iɔ	#144	EGGPLANT (2)	茄	/*giɔ^{2}/
	#218	FEET	腳	/*kiɔk^{7}/
*uɔ	#149	GLUTINOUS RICE	糯米	/*nuɔ^{6} ʔmɐj^{3}/
	#095	WOOD STAKE	木樁	/*muk^{8} tʃuɔŋ^{1}/
	#001	DO ON BEHALF	幫	/*puɔŋ^{1}/

Table 8.2: Examples of Common Yue Nuclei Before Zero and Velar Codas

8.3 Vowel Nuclei Elsewhere

8.3.1 Non-open Vowel Nuclei

Table 8.3 demonstrates the correspondence of the CY non-open vowels before other codas. The <G> notation stands for all grave onsets. These correspondences are much less intuitive at first inspection than the ones discussed above because most of the high vowels have fallen along the nonperipheral track to /ɐ/ according to the Upper Exit Principle. The greatest support for the highly obscured original state reconstructed for Common

CY Nucleus Condition	i /_w	i	y	u /G_	u /_j	u	ɐ /_j	ɐ /_w	ɐ
YS	ɐ	ɐ	ɐ	ɐ	o/y	y	ɐ	a	a/ɐ
LX	ɐ	ɐ	ɐ	ɐ	œ/y	y	ɐ/a	ɐ/a	ɐ
HJ	a	ɐ	ɐ	ɐ	ɔ/u	y	ɐ	a	ɐ
FK	ɔ	ɐ	ɐ	ɐ	œ	y	ɐ	ɔ	ɐ
GZ	ɐ	ɐ	œ	ɐ	œ	y	ɐ	ɐ	ɐ
NH	ɐ	ɐ	œ	ɐ	ɔ	y	ɐ	a/ɐ	ɐ
DG	a	ɔ/ɐ	ø/ɐ	u	ø	u	ɔ	ɔ/a	ɔ/ɐ
FG	ɐ	ɐ	ɐ	ɐ	ø	y	ɐ	ɐ	ɐ
YF	ɐ	ɐ	ɐ	ɐ	u	y	ɐ	ɐ	ɐ/e
XH	æ	æ	æ	æ	u	æ	æ	æ	æ
TS	i	i	u	u	u	u	a	e	e
EP	i	i/ia	u	u	u	u	a	i	ia
KP	ia	i	u	u	u	u	a	a	a
BL	ɐ	ɐ	œ	ɐ	u	u	ɐ	ɐ	ɐ
QZ	ɐ	ɐ	œ	ɐ	u	i	ɐ	ɐ	ɐ
NN_p	ɐ	ɐ	ɐ	ɐ	u	ɔ	ɐ	ɐ	ɐ
BS_p	ɐ	ɐ	ɐ/ə	ɐ/ə	u	ɔ	ɐ	ɐ	ɐ
BY	ə	ɐ/ə	ə	ɐ	œ	u	ɐ/ə	ɐ	ɐ
MS	ɐ	ɐ/ə	ɐ/ə	ɐ	u	u	ɐ/i	ɐ	ɐ
YJ	ɐ/iɛ	ɐ	u	ɐ	u	u	ɐ	ɐ	ɐ
WC	ɐ	ɐ	ɐ	ɐ	u	u	ɐ	ɐ	ɔ/ɐ

Table 8.3: Non-open Vowel Nuclei Elsewhere

Yue comes in the TS, EP, and KP dialects. CY /*i/ did not exit the peripheral in the TS and KP dialects. In the EP dialect, we find it as /ia/. It likely developed from the intermediate stage /iɐ/. In the Sìyì group everywhere, CY /*y/ merged with /*u/ in the the non-zero coda environment. This new /u/ phoneme did not fall either in the TS, EP, and KP dialects.

Characterizing the TS, EP, and KP dialects' correspondences as innovations rather than

retentions is not a persuasive way to explain the situation. First, CY /*i/, although it has fallen, conditioned palatalization of velar onsets in certain environments in the the Fēnglián and Yōngxún dialects, as discussed earlier. One must suppose that this vocalic nucleus originally carried the feature [front]. Similarly, CY /*y/ demonstrates the features front and round in modern correspondences, while CY /*u/ frequently retains its original round feature in the form of secondary labiovelar onsets. One might suppose these features were carried by original medials, but this lacks persuasive power since similar vowel nuclei are not found before velar and open codas. Furthermore, this would suggest loss of the medial for all dialects in the study. Such a proposal bears more sound change assumptions than characterizing the monophthongs in the TS, EP, and KP dialects as cases of retention. If we do not posit these vowels as the beginning of the vowel chain shift, we will have difficulty explaining the motivation other parts of the chain.

In the DG dialect, the intermediate form /ɐ/ further rounded to /ɔ/ before labial codas before labial codas were lost entirely. In the EP dialect, CY /*i/ did not break to /ia/ after palatal onsets. In the BY dialect, /*i/ always centered to /ə/ before the coda /-w/. Before other codas it only fell to /ə/ after palatal and post-alveolar onsets and fell to /ɐ/ elsewhere. This second trend holds true in the MS dialect as well.

In the DG dialect, /*y/ unrounded before stop codas, but stayed round elsewhere. /*y/ evolved into /ə/ in the BS_p, BY, and MS dialects after palatal and post-alveolar onsets, but became /ɐ/ elsewhere

In the Guǎngfǔ and Fēnglián groups, /*u/ tended to front before /*-j/. Additionally,

in the YS, LX, and HJ dialects of the Fēnglián the final /*-uj/ reduced to either /y/ or /u/ after post-alveolar onsets. In the BY dialect, this final always reduced to /œ/.

/*u/ before acute codas is marginal, just appearing once in our data in etymon #071 GRANDSON /*sun[1]/. In the Guǎngfǔ and Fēnglián groups, it evolved the same way /uɔ/ did between acute onsets and codas. They two nuclei merged as /u/, then fronted to /y/. In the Sìyì group, the nucleus /u/ remained mostly unchanged in this environment. In the NN_p and BS_ dialects, it evolved into /ɔ/. Just as the case was with the nucleus /u/ in open syllables in syllables that ended in velar codas in these dialects, it seems /u/ broke to /ɔw/ and then simplified to /ɔ/. These dialects also preserve the original state of the nucleus.

The NH dialect exhibits both /ɐ/ and /a/ as correspondences of CY /*ɐ/ before the coda /-w/. This is likely due to dialect mixture. We suppose a weak change of /ɐ/ exiting the nonperipheral track as /a/ before /-w/ that was partially obscured due to Guǎngfǔ influence.

The DG dialect rounded /ɐ/ to /ɔ/ before labial codas, as discussed before, as well as before the coda /-w/. The same rounded also occurred before labial codas in the WC dialect.

CY /*ɐ/ evolved into /e/ before labial codas in the TS dialect. It evolved /i/ before the coda /w/ in the EP dialect, yielding the diphthong /-ij/.

Nucleus	a	a	ɔ	ɔ	ɔ	ɔ	ia	iɔ	uɔ	uɔ	uɔ
Cond.	/K_{P,T}		/_j	/_w	/_P	/_T			/G_	/_j	/_T
YS	e	a	o	o	a	o	i/e	y	o	o/u	y
LX	ɛ	a	œ	o	a	o	i/ɛ	y	u	u/œ	y
HJ	ɛ	a	ɔ	ɔ	a	ɔ	i	y	u	œ	y
FK	a	a	ɔ	ɔ	ɔ	ɔ	i/e	y	ɐ	œ	y
GZ	a	a	ɔ	u	ɐ	ɔ	i	y	u	œ	y
NH	ɛ/i	a	ɔ	ɔ	u	ɔ	i	y	u	y	y
DG	ɛ	a	u	ɔ	a	u	i	ø	u	u	ø
FG	a	a	ø	a	a	ɔ	i	y	ø/u	ø	y
YF	a	a	u	u	ɐ	u	i	y	u	u	y
XH	a	a	u	u	æ/a	u	i	i	u	u	u
TS	a	a	ɔ	a	a	ɔ	ia/e	u	ɔ	u	ɔ
EP	a	a	ua	a	a	ua	i	i	(ɔ)/ua	u	ua
KP	a	a	ɔ	ɔ	a	ua	i	i	ɔ/ua	u	ua
BL	a	a	o	ɐ	ɐ	o	i	(j)o/i	u	u	y
QZ	a	a	ɔ	ɐ	a	ɔ	i	i	u	u	i
NN_p	a/ɛ	a	a	a	a	a	i	(w)i	u	ɔ	u
BS_p	a	a	a	a/ɐ	a	a	i	(w)i	u	ɔ	u
BY	a	a	a	œ	œ	œ	i	(j)u	u	u	u
MS	a	a	ɔ	ə	ə	a	i	(j)u	u	u	u
YJ	a	a	ɔ	u	ɐ	ɔ	i	i	u	u	u
WC	a	a	ɔ	ɔ	ɔ	ɔ	i	(w)i/u	u	u	u

Table 8.4: Open Vowel Nuclei Elsewhere

8.3.2 Open Vowel Nuclei

For the most part, CY /*a/ remains unchanged before the remaining codas in the dialects. However, after velar onsets before the nasal and stop codas /*a/ shifted forward to /ɛ/ in most of the Fēnglián group dialects as well as in the NH dialect. [1] This is a weak sound change in the NH dialect. It is observable in most of our etyma involving /a/ in these

[1]The NH dialect has /i/ in this position before CY dental codas that have now shifted to velars .

positions, but the correspondences tend to be erratic in some words. This weak sound change is yet another expression of expanding Guǎngfǔ influence in the Northern Pearl River Delta region. As an example of this weak change, in etymon #203 EYE 眼 /*ʔŋan^{3}/, the NH dialect all exhibits /ɛ/ as the modern nucleus, but in etymon #063 WILD GOOSE 雁 /*ŋan^{6}/ it exhibits /a/.

In the DG dialect, CY /*a/ shifted to /ɛ/ before dental codas, but remained /a/ before labial codas. Although /*a/ has not changed in the Yōngxún group, in the BY dialect velar onsets palatalize in this position and the velar nasal onset palatalizes in all the inland Yōngxún dialects.

In Table 8.4 labial codas are marked with <P> and dental codas are marked as <T> as a shorthand.

CY /*ɔ/ remained largely stable before the coda /*-j/ and dental codas, with the exception of the Yōngxún dialects where /*ɔ/ has fallen to /a/ generally. Before the coda /*-w/, /*ɔ/ changed the most. The best represented changes are raising to /uw/ [ou], monophthongization to /ɔ/ and /u/, and lowering to /aw/. Before labial codas, /* ɔ/ mainly centered to /ɐ/, raised to /u/, and fell to /a/ among the dialects. The XH dialect demonstrates both the centering and falling trends. Because the correspondences for /*ɔ/ before labial codas are marginal, it is difficult to ascertain which correspondences are autochthonous to those dialects and which may be the product of dialect mixture.

In most dialects, CY /*ia/ has monopthongized and raised to /i/. In the TS dialect, it remains unchanged before labial codas and has monophthongized to /e/ before dental

codas.

In the Fēnglián group, most dialects preserve a lower vowel /e/ or /ɛ/ as a reflex of this nucleus after dental onsets. The FK dialect seems to also intermittently reflect this lower vowel after alveolar sibilants as well.

As mentioned previously, in the Zēngchéng 增城 dialect CY /*ia/ commonly coincides with modern /ɛ/ before dental codas which have since backed to velars. This change predominately occurred after onsets other than palatal and post-alveolar onsets.

The vowel correspondences for /*iɔ/ suggest that it raised to /y/ in virtually all dialects. However, it cooccurs with tone 9 in many dialects before the coda /*-t/ so we reconstruct it as a complex nucleus ending in an open vowel by analogy. In the Zēngchéng dialect, this nucleus commonly correspondences to modern /œ/ before dental codas, which have since backed to velar codas. This correspondence is best represented after onsets other than palatal and post-alveolar onsets.

CY /*uɔ/ monophthonized to /*o/ and then raised to /u/ in most dialects. In the Sìyì group, it merged with /ɔ/ in the TS, EP, and KP dialects, except before the coda /-j/ where it raised to /u/. After grave initials, the new /u/ mainly stayed in place. After acute initials, new /u/ shifted commonly changed to further front to /y/. This lowered to /œ/ before /-j/ in some dialects, causing a merger with CY /*-uj/ in many dialects. However, due to further changes of CY /*-uj/ involving coda loss in NH dialect and low vowel reflexes in the Yōngxún group, the two finals remain distinct in the common system.

Nucleus	Examples			
*i	#098	DUST	塵	/*dʒin^{2}/
	#120	NINE	九	/kiw^{3}/
*y	#081	BOTTLE	樽	/*tsyn^{1}/
	#127	CHESTNUT	栗	/*lyt^{8}/
*u	#225	PUPIL		/*ʔŋan^{3} wut^{8}/
	#071	GRANDSON	孫	/*sun^{1}/
	#220	MOUTH	嘴	/*tsuj^{3}/
*ɐ	#103	SILVER	銀	/*ŋɐn^{2}/
	#214	HEAD	頭	/*dɐw^{2}/
	#101	MUD	泥	/*nɐj^{2}/
*a	#114	THREE	三	/*sam^{1}/
	#119	EIGHT	八	/*pat^{7}/
	#023	CRAB	蟹	/*xaj^{3}/
*ɔ	#184	DRY	乾	/*kɔn^{1}/
	#162	PEACH	桃	/*dɔw^{2}/
	#174	SWEETPEEL TANGERINE	柑	/*kɔm^{1}/
*ia	#123	THOUSAND	千	/*ts^{h}ian^{1}/
	#017	BUTTERFLY	蝴蝶	/*wu^{2} diap^{8}/
	#134	BANANA	蕉	/*tsiaw^{1}/
*iɔ	#100	MOON	月	/*ŋiɔt^{8}/
	#200	COWLICK (HAIR)	旋	/*dziɔn^{6}/
*uɔ	#146	GARLIC	蒜	/*suɔn^{1}/
	#086	DOOR	門	/*muɔn^{2}/
	#201	DEAF	耳背	/*ʔɲi^{3} buɔj^{6}/

Table 8.5: Examples of Common Yue Nuclei Elsewhere

8.4 Marginal Vowel Nuclei

#002 EAT, #006 PICK UP WITH CHOPSTICKS, #018 CAT, #088 PINCERS, #194 BRAID, #195 BREASTS, and #211 GRASP IN ARMPIT all exhibit erratic correspondences of vowel nuclei. #006 PICK UP WITH CHOPSTICKS, #018 CAT, and #088 PINCERS seem to be the most stable and have cases of the open front vowel /ɛ/ in finals with closure at the

front of the mouth in the Yōngxún dialects and occasionally elsewhere, something that is normally phonologically prohibited. For the nonce, we will set up a nucleus /*ɛ/ to handle these words' correspondences, but need more data to be sure that this phoneme indeed contrasts with /*ia/ in the common system.

Chapter 9

CONCLUSION

We hope that this new comprehensive reconstruction of historical Yuè phonology and contextualization of *koiné* influence in Yuè dialects will aid other investigators in comparison of Yuè dialects with other languages of the region as well facilitate a more sophisticated discussion of lexical layers and dialect classification in the modern Lǐngnán region. A few avenues for future study come to mind.

The next step in refining the CY common system would be to attempt to seriate character readings into layers based on the correspondences we have proposed here. With this seriation in hand, we can better understand the nature of the *koiné* loans and put us in a position where we can begin to compare Yuè character readings to other varieties of Chinese, like Hakka and Shē.

To do this, it might be desirable to first obtain more cognate sets first. However, one thing that became apparent during this study was that many words themselves, like #051 PIG, #082 CHOPPING BOARD, #091 STEAMED RICE, and #132 BAMBOO already exhibited mixed correspondences that suggest lexical layering. #091 STEAMED RICE mixes bilabial stop and labiodental onsets. The others mix dental stop and post-alveolar sibilant onsets. Clearly, focusing on words instead of character readings does not com-

pletely avoid the issue of lexical layering. At this stage, it is difficult to tell how much of this might be due to the way vocabulary was elicited since none of the surveys describe their process. However, it does seem clear at least some of this variation, if not the vast majority of it, is due to lexical replacement spurred on by waves of prestigious regional standards entering the area. Nonetheless, I would still urge other investigators to look at lexical documentation of the dialects they are comparing. Knowing what words people actually use for certain concepts can help in deciding whether or not a character reading is of later *koiné* origin, i.e., "artificial." This kind of investigation is also the only way to recover words without standard orthography. It is needed to recover early non-sinitic loanwords and sinitic words from substrates that do not have standard orthography associated with them. Examples of these types of words are #168 POMELO and #200 COWLICK, respectively.

With a refined common system and greater inventory of reconstructed forms, one could begin to tackle the question of the relationship of Yuè dialects with other languages in the area like Coblin's Early Southern Highlands Chinese and the various Zhuang languages of the region. Yuè dialects are rich in non-sinitic loans. Some of these, like the word for "spider" have irregular aspiration in low register tones, suggesting a late date of borrowing. It would be worth pursuing a chronology of when these words entered the language to better understand the contact situation.

Hopefully, the common system presented here has successfully built upon the previous diachronic studies of Yuè phonology so that it might be used in further studies on "South

Central" Chinese languages.

BIBLIOGRAPHY

Akitani, Hiroyuki 秋谷裕幸 and Handel, Zev, 2012. Lìshǐ bǐjiàofǎ hé céngcì fēnxīfǎ 歷史比較法和層次分析法. *Yǔyánxué lùncóng*, 45, pp.277–335.

Ball, J. Dyer, 1889. The san-wúí dialect. *The china review*, 18 (3), pp.178–195.

— 1890. The tung-kwún dialect. *The china review*, 18 (5), pp.284–299.

— 1896. The höng shán or macao dialect. *The china review*, 22 (2), pp.501–531.

— 1900a. The shun tak dialect. *The china review*, 25 (2), pp.57–69.

— 1900b. The shun tak dialect (cont.) *The china review*, 25 (3), pp.121–140.

Bāo Zhìmíng 包智明, 1996. The syllable in chinese. *Journal of chinese linguistics*, 24(2), pp.312–354.

Bauer, Robert S and Benedict, Paul K, 2011. *Modern cantonese phonology*. Vol. 102. Walter de Gruyter.

Baxter, William H, 1992. *A handbook of old chinese phonology*. Walter de Gruyter.

Bielenstein, Hans, 1947. The census of china during the period 2–742a.d. *Bmfea*, 19, pp.125–163.

Chao Y.R. 趙元任, 1948a. Táishān yǔliào 臺山語料. *Zhōngyāng yánjiūyuàn lìshǐyǔyán yánjiūsuǒ jíkān*, 23, pp.25–76.

— 1948b. Zhōngshān fāngyán 中山方言. *Zhōngyāng yánjiūyuàn lìshǐyǔyán yánjiūsuǒ jíkān*, 20, pp.49–73.

Chen, Matthew Y. and Wang, William S.Y., 1975. Sound change: actuation and implementation. *Language*, 51, pp.255–281.

Chén Zhōngmǐn 陳忠敏, 2013. Yě tán lìshǐ bǐjiàofǎ yǔ lìshǐ céngcì fēnxīfǎ 也談歷史比較法與歷史層次分析法. *Yǔyánxué lùncóng*, 47, pp.35–69.

Coblin, W. South, 2015. *A study of comparative gàn*, Language and Linguistics Monograph Series 58. Institute of Linguistics, Academia Sinica.

— 2018. Neo-hakka, paleo-hakka, and early southern highlands chinese. *Yǔyán yánjiū jíkān*, 21, pp.175–238.

— n.d.(a). *Early south highlands chinese - a comparative historical study*. Unpublished Manuscript. Unpublished.

— n.d.(b). *The reconstruction of common shē and common hakka shē*. Unpublished Manuscript. Unpublished.

DeSousa, Hilário, 2017. The pínghuà 平話 dialects. *Encyclopedia of chinese language and linguistics*. Brill.

Don, Alexander, 1882. The llin-nen variation of cantonese. *The china review*, 11, pp.236–247.

— 1884. The llin-nen variation of cantonese (cont.) *The china review*, 12, pp.474–481.

Duanmu, San, 2007. *The phonology of standard chinese*. OUP Oxford.

Gě Jiànxióng 葛劍雄, 2002. *Zhōngguó rénkǒushǐ* 中國人口史. Ed. by Jiànxióng Gě 葛劍雄. Vol. 1. Shànghǎi: Fùdàn dàxué chūbǎnshè.

Gě, Jiànxióng 葛劍雄, Cáo, Shùjī 曹樹基, and Wú, Sōngdì 吳松弟, 1997. *Zhōngguó yímínshǐ* 中國移民史. 6 vols. Fújiàn: Fújiàn rénmín chūbǎnshè.

Golston, Chris and Kehrein, Wolfgang, 2015. A prosodic theory of vocalic contrasts. In: *The segment in phonetics and phonology* [Online]. John Wiley & Sons, Ltd. Chap. 4, pp.65–102. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118555491.ch4. Available from: https://doi.org/10.1002/9781118555491.ch4.

Handel, Zev, 2015. Huíyīng chén zhōngmǐn duìyú “lìshǐbǐjiàofǎ hé céngcì fēnxīfǎ” de huíyīng 回應陳忠敏對於《歷史比較法何層次分析法》的回應. *Yǔyánxué lùncóng*, 51, pp.331–339.

Heeringa, Wilbert Jan, 2004. *Measuring dialect pronunciation differences using levenshtein distance*. PhD thesis. University Library Groningen][Host].

Huang, Karen, 2009. A reconstruction of proto-yue vowels.

Jeszenszky, Péter, Hikosaka, Yoshinobu, Imamura, Satoshi, and Yano, Keiji, 2019. Japanese lexical variation explained by spatial contact patterns. *Isprs international journal of geo-information*, 8(9), p.400.

Kehrein, Wolfgang, 2013. *Phonological representation and phonetic phasing: affricates and laryngeals*. Publication Title: Phonological Representation and Phonetic Phasing. De Gruyter.

Kehrein, Wolfgang and Golston, Chris, 2004. A prosodic theory of laryngeal contrasts. *Phonology* [Online], 21(3). Publisher: Cambridge University Press, pp.325–357. Available from: https://www.jstor.org/stable/4615515 [Accessed June 4, 2020].

Kleiweg, Peter, 2007a. *How to determine the best clustering of an area* [Online]. Available from: http://www.let.rug.nl/~kleiweg/kaarten/MDS-clusters.html.en.

— 2007b. *How to map difference between geographic areas* [Online]. Available from: http://www.let.rug.nl/~kleiweg/kaarten/Afstanden.html.en#minclust.

— 2007c. *Tutorial 7, example: germany* [Online]. Available from: http://www.let.rug.nl/~kleiweg/L04/Tutorial/t07.html.en.

Kleiweg, Peter, Nerbonne, John, and Bosveld, Leonie, 2004. Geographic projection of cluster composites. *International conference on theory and application of diagrams*. Springer, pp.392–394.

Kruskal, Joseph B, 1964. Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29(2), pp.115–129.

Kwok Bit-chee 郭必之, 2006. Yíduàn wēiguān de yuèyǔ shǐ: zhūjiāng sānjiǎozhōu bāzhǒng yuèyǔ fāngyán de fēnqún 一段微觀的粵語史：珠江三角洲八種粵語的分群. In: *Shān gāo shuǐ cháng: dīng bāngxīn xiānshēng qīzhìshòuqìng lùnwénjí* 山高水長：丁邦新先生七秩壽慶論文集. Taipei: Institute of Linguistics, Academia Sinica, pp.701–739.

Labov, William, 1981. Resolving the neogrammarian controversy. *Language*, 57, pp.267–308.

— 1989. Exact description of the speech community: short a in philadelphia. *Language change and variation*, pp.1–57.

— 1994. *Principles of linguistic change: internal factors*. Oxford: Basil Blackwell.

— 2007. Transmission and diffusion. *Language*, 83(2), pp.344–387.

— 2010. *Principles of linguistic change, volume 3: cognitive and cultural factors*. John Wiley & Sons.

— 2020. The regularity of regular sound change. *Language*, 96(1), pp.42–59.

Labov, William, Ash, Sharon, and Boberg, Charles, 2008. *The atlas of north american english: phonetics, phonology and sound change*. Walter de Gruyter, pp.16–20.

Levenshtein, Vladimir I, 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics doklady*. Vol. 10, 8, pp.707–710.

Li Fang-Kuei 李方桂, 1973. Languages and dialects of china. *Journal of chinese linguistics*, 1, pp.1–13.

Lǐ Jiàn 李健, 2014. *Wú huà yuèyǔ yánjiū* 吳化粵語研究, Zhānjiāng shīfàndàxué zhōngguó yǔyánxué xuékē xīn shìyě xuéshù wéncóng. Zhōngguó shèhuìkēxué chūbǎnshè.

Lǐ Liánjìn 李連進, 2000. *Pínghuà yīnyùn yánjiū* 平話音韻研究. Guǎngxī rénmín chūbǎnshè.

— 2005. Gōulòu piān de fāngyán guīshǔ 勾漏片的方言歸屬. *Mínzú yǔwén*. 1, pp.34–41.

Liáng Jīnróng 梁金榮, 1997. *Guìnán pínghuà yǔyīn yánjiū* 桂北平話語音研究. PhD thesis. Jìnán dàxué.

Light, Timothy, 1977. The cantonese final: an exercise in indigenous analysis. *Journal of chinese linguistics*, 5(1), pp.75–102.

Lin, Yen-Hwei, 1989. *Autosegmental treatment of segmental processes in chinese phonology*. PhD thesis. The University of Texas at Austin.

Maddieson, Ian, 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.

Malkiel, Yakov, 1962. Weak phonetic change, spontaneous sound shift, lexical contamination. *Lingua*, 11, pp.263–275.

Malkiel, Yakov, 1964. Some diachronic implications of fluid speech communities. *American anthropologist*, 66(6), pp.177–186.

McCoy, John, 1966. *Szeyap data for a first approximation of proto-cantonese*. Cornell University.

— 1986. Modern suprasegmental evidence for the consonant clusters in proto-yue. In: *Contributions to sino-tibetan*. Ed. by John McCoy and T. Light. Leiden: E.J. Brill, pp.367–374.

Morrison, Robert, 1828. *A vocabulary of the canton dialect*. London: Ganesha.

Nakanishi Hiroki 中西裕樹, 2010. Lùn shēhuà de guīshǔ 論畲話的歸屬. In: *Diachronic change and language contact–dialect in south china*. Ed. by Hong-nin Cheung and Song Hing Chang, Journal of Chinese Linguistics Monograph Series. Institute of Linguistics, Academia Sinica.

Nerbonne, John and Heeringa, Wilbert, 1997. Measuring dialect distance phonetically. *Computational phonology: third meeting of the acl special interest group in computational phonology*.

Nerbonne, John, Heeringa, Wilbert, Van den Hout, Erik, Van der Kooi, Peter, Otten, Simone, Van de Vis, Willem, et al., 1996. Phonetic distance between dutch dialects. *Clin vi: proceedings of the sixth clin meeting*, pp.185–202.

Nerbonne, John and Kleiweg, Peter, 2003. Lexical distance in lamsas. *Computers and the humanities*, 37(3), pp.339–357.

— 2007. Toward a dialectological yardstick. *Journal of quantitative linguistics*, 14(2-3), pp.148–166.

Nerbonne, John, Kleiweg, Peter, Heeringa, Wilbert, and Manni, Franz, 2008. Projecting dialect distances to geography: bootstrap clustering vs. noisy clustering. *Data analysis, machine learning and applications*. Springer, pp.647–654.

Nerbonne, John and Kretzschmar, William, 2003. Introducing computational techniques in dialectometry. *Computers and the humanities*, 37(3), pp.245–255.

Norman, Jerry, 1988. *Chinese*, Cambridge Language Surveys. Cambridge University Press.

— 2006. Common dialectal chinese. In: David Branner, ed. *The chinese rime tables: linguistic philosophy and historical-comparative phonology*. John Benjamins Publishing, pp.233–254.

Ohala, John J, Hombert, Jean-Marie, and Ewan, William G, 1979. Phonetic explanations for the development of tones. *Language*, pp.37–58.

Péng Xiǎochuān 彭小川, 2005. *Yuèyǔ lùngǎo* 粵語論稿. Guǎngzhōu: Jìnān dàxué chūbǎnshè.

Qín Yuǎnxióng 覃遠雄, 2000. *Guìnán pínghuà yánjiū* 桂南平話研究. PhD thesis. Jìnán dàxué.

— 2007. Pínghuà hé tǔhuà 平話和土話. *Fāngyán*. 2, pp.177–189.

Saunders, C.J., 1896. The tungkwun dialect of cantonese. *The china review*, 22 (1), pp.465–476.

Smith, Thomas and Guild, John, 1931. The cie colorimetric standards and their use. *Transactions of the optical society*, 33(3), p.73.

Ting Pang-hsin 丁邦新, 1982. Hànyǔ fāngyán fēnqū de tiáojiàn 漢語方言的分區條件. *Qīnghuá xuébào*, 14, pp.257–273.

— 1996. Píng “zhōngguó yǔyán dìtújí” 評《中國語言地圖集》. *Guójì zhōngguó yǔyánxué pínglùn*, 1.

Tsuji, Nobuhisa, 1980. *Comparative phonology of guangxi yue dialects*. Kazama Shobu Publisng Co.

Valls, Esteve, Nerbonne, John, Prokic, Jelena, Wieling, Martijn, Clue, Esteve, and Lloret, Maria Rosa, 2012. Applying the levenshtein distance to catalan dialects: a brief comparison of two dialectometric approaches. *Verba: anuario galego de filoloxía*, 39.

Wang, William S.Y., 1977. *The lexicon in phonological change*. The Hague: Mouton.

Wáng Fútáng 王福堂, 2001. Pínghuà xiāngnán tǔhuà hé yuèběi tǔhuà de guīshǔ 平話、湘南土話粵北土話的歸屬. *Fāngyán*. 2, pp.107–118.

Wáng Lì 王力 and Qián, Sōngshēng 錢淞生, 1950. Táishān fāngyīn 台山方音. *Lǐngnán xuébào*, 10 (1), pp.67–104.

Wéi Shùguān 韋樹關, 1996. Shìlùn pínghuà zài hànyǔ fāngyán zhōng de dìwèi 試論平話在漢語方言中的地位. *Yǔyán yánjiū*. 2, pp.95–101.

Wieling, Martijn and Nerbonne, John, 2011. Measuring linguistic variation commensurably. *Dialectologia: revista electrònica*, pp.141–162.

Wǔ Wēi 伍巍, 2001. Lùn guìnán pínghuà de yuèyǔ xìshǔ 論桂南平話的粵語係屬. *Fāngyán*. 2, pp.133–141.

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

Xiǎn Wéntíng 冼文婷, 2016. *Guǎngdōng yángjiānghuà yánjiū* 廣東陽江話研究. Master's thesis. Guǎngxī dàxué.

Xiè, Jiànyóu 謝建猷, 2001. *Guǎngxī pínghuà yánjiū* 廣西平話研究. PhD thesis. Zhōngguó shèhuìkēxué yuàn.

— 2007. *Guǎngxī hànyǔ fāngyán yánjiū* 廣西漢語方言研究. 2 vols. Guǎngxī rénmín chūbǎnshè.

Xióng Zhènghuī 熊正輝, 1987. Guǎngdōng fāngyán de fēnqū 廣東方言的分區. *Fāngyán*. 3, pp.161–165.

Yáng Huàndiǎn 楊煥典, Liáng, Zhènshì 梁振仕, Lǐ, Pǔyīng 李譜英, and Liú, Cūnhàn 劉村漢, 1985. Guǎngxī de hànyǔ fāngyán (gǎo) 廣西的漢語方言（稿）. *Fāngyán*. 3, pp.181–190.

Yip, Moira, 1997. Consonant-vowel interaction in cantonese. In: *Studies in chinese phonology*. Ed. by Jialing Wang and Norval Smith. New York: Walter de Gruyter, pp.251–270.

Yuán Jiāhuá 袁家驊 et al., 1983. *Hànyǔ fāngyán gàiyào* 漢語方言概要. Běijīng: Yǔwén chūbǎnshè.

Yue-Hashimoto, Anne, 1972. *Phonology of cantonese*. Vol. 1. Cambridge University Press.

— 1979. *The tengxian dialect of chinese: its phonology, lexicon, and texts with grammatical notes*. Vol. 3, Computational Analysis of Asian and African Languages Monograph Series. Tōkyō: Institute for the Study of Language, Cultures of Asian, and Africa.

— 1988. A preliminary investigation into the subclassification problem of the yue dialects. *Ajia-afurikago no keisū kenkyū*, 30, pp.7–42.

— 1991. Yuèyǔ fāngyán fēnqū wèntí chūtān 粵語方言分區問題初探. *Fāngyán*, 1991 (3), pp.164–181.

Yue-Hashimoto, Anne, 2000. Yuèyīn gòunǐ zhī yī: shēngdiào 粵音構擬之一：聲調. *Zhōngwén xuékān*. 2, pp.227–241.

— 2002. Development of the stop endings in the yue dialects. In: *Nánběi shìfēi: hànyǔ fāngyán de chāyì yǔ biànhuà*. Ed. by Dà'ān Hé, Title in English on p. [4] of cover:Dialect variations in Chinese. Táiběi: Zhōngyāng yánjiūyuàn yǔyánxué yánjiūsuǒ, pp.217–245.

— 2005. *The dancun dialect of taishan*. Language Information Sciences Research Centre, City University of Hong Kong.

— 2006. Yuèyīn gòunǐ zhī èr: shēngmǔ 粵音構擬之二：聲母 [Reconstructing Yue Chinese Part Two: Initials]. In: *Shān gāo shuǐ cháng: dīng bāngxīn xiānshēng qīzhìshòuqìng lùnwénjí* 山高水長：丁邦新先生七秩壽慶論文集. Taipei: Institute of Linguistics, Academia Sinica, pp.75–170.

Zhān Bóhuì 詹伯慧, 1981. *Xiàndài hànyǔ fāngyán* 現代漢語方言. Húběi rénmín chūbǎnshè.

Zhān Bóhuì 詹伯慧 et al., 2000. *Guǎngdōng yuè fāngyán gàiyào* 廣東粵方言概要. Jìnán dàxué chūbǎnshè.

Zhān Bóhuì 詹伯慧, 2001. Guǎngxī "pínghuà" wèntí chúyì 廣西"平話"問題芻議. *Yǔyán yánjiū*. 2, pp.92–99.

Zhān Bóhuì 詹伯慧, Cheung, Yat-shing 張日昇, et al., 1987a. *Zhūjiāng sānjiǎozhōu fāngyán cíhuì duìzhào* 珠江三角洲方言詞彙對照. Guǎngdōng Rénmīn Chūbǎnshè.

Zhān Bóhuì 詹伯慧, Cheung, Yat-shing 張日昇, et al., 1987b. *Zhūjiāng sānjiǎozhōu fāngyán zìyīn duìzhào* 珠江三角洲方言字音對照. Guǎngdōng Rénmīn Chūbǎnshè.

— 1994. *Yuèběi shí xiànshì yuè fāngyán diàochá bàogào* 粵北十縣市粵方言調查報告. Jìnán Dàxué Chūbǎnshè.

— 1998. *Yuèxī shí xiànshì yuè fāngyán diàochá bàogào* 粵西十縣市粵方言調查報告. Jìnán Dàxué Chūbǎnshè.

Zhān Bóhuì 詹伯慧, Cuī, Shūhuì 崔淑慧, Liú, Xīnzhōng 劉新中, and Yáng, Wèi 楊蔚, 2003. Guānyú guǎngxī “pínghuà” de guīshǔ wèntí 關於廣西“平話”的歸屬問題. *Yǔyán yánjiū*. 3, pp.47–52.

Zhāng Jūnrú 張均如, 1982. Guǎngxī nánbù dìqū zhuàngyǔ zhōng de lǎo jiècí yuán yú hànyǔ gǔ “pínghuà” kǎo 廣西南部地區中的老借詞源於漢語古“平話”考. *Yǔyán yánjiū*. 1, pp.197–219.

Appendix A

MINIMAL FEATURE ENCODING OF PHONETIC SEGMENTS

Segment	Encoded Features			
i	Open0	High1	Front1	Round0
y	Open0	High1	Front1	Round1
e	Open0	High0	Front1	Round0
ø	Open0	High0	Front1	Round1
ɛ	Open1	High0	Front1	Round0
ɛ	Open0	High0	Front1	Round0
œ	Open1	High0	Front1	Round1
æ	Open1	High0	Front1	Round0
a	Open1	High0	Front0	Round0
ɑ	Open1	High0	Front0	Round0
ɨ	Open0	High1	Front0	Round0
ɵ	Open0	High0	Front0	Round1
ə	Open0	High0	Front0	Round0
ɐ	Open0	High0	Front0	Round0
ɯ	Open0	High1	Front0	Round0

Segment	Encoded Features			
u	Open0	High1	Front0	Round1
ʏ	Open0	High0	Front0	Round0
o	Open0	High0	Front0	Round1
ʌ	Open1	High0	Front0	Round0
ɔ	Open1	High0	Front0	Round1
ɑ	Open1	High0	Front0	Round0
ɥ	Open0	High1	Front0	Round1
ɭ	Open0	High1	Front0	Round0
p	Labial	Stop	Voice0	
b	Labial	Stop	Voice1	
m	Labial	Nasal	Voice1	
w	Labiovelar	Approximant	Voice1	
f	Labial	Fricative	Voice0	
v	Labial	Fricative	Voice1	
t	Dental	Stop	Voice0	
d	Dental	Stop	Voice1	
n	Dental	Nasal	Voice1	
s	AlveolarPlus	Fricative	Voice0	
θ	Dental	Fricative	Voice0	

Segment	Encoded Features			
z	AlveolarPlus	Fricative	Voice1	
ʃ	AlveolarPlus	Fricative	Voice0	
tʃ	AlveolarPlus	Stop	Voice0	
dʒ	AlveolarPlus	Stop	Voice1	
ɬ	Dental	Fricative	Voice0	
l	Dental	Approximant	Voice1	
ts	AlveolarPlus	Stop	Voice0	
dz	AlveolarPlus	Stop	Voice1	
ɲ	AlveolarPlus	Nasal	Voice1	
n̠	AlveolarPlus	Nasal	Voice1	
j	AlveolarPlus	Approximant	Voice1	
k	Velar	Stop	Voice0	
g	Velar	Stop	Voice1	
ŋ	Velar	Nasal	Voice1	
ʔ	Glottal	Stop	Voice0	
h	Glottal	Fricative	Voice0	SpreadGlottis
ɕ	AlveolarPlus	Fricative	Voice0	
ʑ	AlveolarPlus	Fricative	Voice1	
tɕ	AlveolarPlus	Stop	Voice0	

Segment	Encoded Features			
dʐ	AlveolarPlus	Stop	Voice1	
ʂ	AlveolarPlus	Fricative	Voice0	
ʐ	AlveolarPlus	Fricative	Voice1	
tʂ	AlveolarPlus	Stop	Voice0	
dʐ	AlveolarPlus	Stop	Voice1	
ʈ	Dental	Stop	Voice0	
ɖ	Dental	Stop	Voice1	
ᵐb	Labial	Nasal	Voice1	
ⁿd	Dental	Nasal	Voice1	
ᵑg	Velar	Nasal	Voice1	
¹	LowA1	LowB1	HighA0	
²	LowA1	LowB0	HighA0	
³	LowA0	LowB0	HighA0	
⁴	LowA0	LowB1	HighA1	
⁵	LowA0	LowB0	HighA1	
ʷ	Open0	High1	Front0	Round1
ʰ	SpreadGlottis			
˜	Nasal			

Appendix B

LEXICAL ITEMS COMPARED

All 77 lexical items used in the measurement of the average lexical distance between sites are listed below. Sources of the indexes are given according to Table 1.1.

Table B.1

Item	[1]; [6]	[2]; [3]	[4]	Chinese Gloss	English Gloss
1	5	2	2	月亮	moon
2	15	13	13	閃電	flash lightning
3	20	18	18	下雨	to rain
4	27	26	26	虹	rainbow
5	53	53	44	涼水	cold water
6	101	104	83	明天	tomorrow
7	104	107	86	昨天	yesterday
8	181	169	141	黃豆	soybean
9	236	223	187	柚子	pomelo
10	270	261	215	公狗	male dog
11	271	262	216	母狗	female dog
12	286	277	231	(下) 蛋	egg
13	286	277	231	下 (蛋)	lay (egg)
14	311	295	249	翅膀	wing

Table B.1

Item	[1]; [6]	[2]; [3]	[4]	Chinese Gloss	English Gloss
15	314	299	253	麻雀	sparrow
16	324	309	262	蝙蝠	bat
17	326	311	264	蜘蛛	spider
18	330	315	268	蒼蠅	fly
19	345	330	283	蜻蜓	dragonfly
20	359	340	292	蝦	shrimp
21	367	348	300	青蛙	frog
22	369	350	302	蟾蜍	toad
23	383	362	311	門檻	door jam
24	567	525	462	鼻子	nose
25	574	532	469	舌頭	tongue
26	586	543	480	肩膀	shoulder
27	603	560	497	乳房	breasts
28	453	429	368	傘	umbrella
29	737	682	614	鹽	salt
30	767	730	657	吃早飯	eat (breakfast)
31	766	729	656	餓了	hungry
32	807	726	653	玩耍	play
33	808	727	654	逛	stroll
34	827	794	715	睡了	sleep
35	898	819	740	乖	well-behaved
36	963	889	803	欠	owe

Table B.1

Item	[1]; [6]	[2]; [3]	[4]	Chinese Gloss	English Gloss
37	1056	604	540	看	look
38	1143	1029	934	找	search for
39	1151	1038	942	害怕	be afraid
40	1193	1124	1028	稀	not viscous
41	1252	1078	982	誰	who
42	1254	1080	984	我	1sg
43	1265	1086	990	我們	(1)pl
44	1274	1097	1001	這個	this
45	1275	1098	1002	那個	that
46	1285			怎樣 (做)	how (to do)
47	1316			全部	all
48	1320			也	also
49	1321	1173	1076	被 (狗咬了)	pass.
50	1327	1179	1082	在	loc.
51	1332	1184		和	and
52	1345	1197	1097	一匹馬	horse cntr
53	1361	1209	1109	一棵樹	tree cntr
54	1378	1224	1122	一趟	instance cntr
55	436	415	358	瓶子	bottle
56	438	417	360	湯匙	soup spoon
57	439	418	361	筷子	chopsticks
58	494	470	408	兒子	son

Table B.1

Item	[1]; [6]	[2]; [3]	[4]	Chinese Gloss	English Gloss
59	1390	1235	1133	(幾) 個	general mw
60	173	161	133	稻子	growing grain
61	336	321	274	蟋蟀	cricket
62	711	655	589	剩飯	leftover (rice)
63	711	655	589	(剩) 飯	steamed rice
64	747	695	623	(雞) 胗	(chicken) gizzard
65	771	734	661	盛飯	serve up rice
66	819	786	707	乘 (涼)	relax in a cool place
67	821	788	709	烤 (火)	keep a fire (for warmth)
68	854			欺負	bully
69	934	862	778	倒閉	go out of business
70	1000	455	393	乞丐	beggar
71	1002	457	395	小偷	thief
72	1187	1118	1022	結實	sturdy
73	1189	1120	1023	骯髒	dirty
74	1206	1137	1041	頑皮	cheeky
75	1346	1196	1096	(一) 把 (刀)	knife mw
76	1121	1005	911	叫	shout
77	1125	1009	915	扔	throw

Appendix C

COMMON YUE RECONSTRUCTED FORMS

The following tables list all comparanda and reconstructed forms in this study. Dialect forms in parentheses are not cognate with the reconstructed form. Forms preceded by an exclamanation mark demonstrate irregular correspondence. Brackets jn reconstructed forms indicate uncertainty due to few correspondence sets. Forms are given in alphabetical order for each lexical domain. The index ranges of the lexical domains of the forms are as follows:

- Actions 1–11
- Animals 12–64
- Kinship 65–79
- Miscellaneous 80-95
- Nature 96–111
- Numerals 112–127
- Plants 128–178
- Pronouns 179–181
- States 182–188
- The Body 189–240

Gloss	do on behalf	eat	know	look	owe
Index	1	2	3	4	5
Orth	幫	喫	知	睇	爭
CY	*puɔŋ^{1}	*x[ɛ][k]^{7}	*ti^{1}	*tʰɐj^{3}	*tʃaŋ^{1}
YS	pɔŋ$^{1-52}$	hik^{-33}	tɐj$^{1-52}$	tʰaj^{2}	tseŋ$^{1-52}$
LX	pœŋ^{1}	hik^{8}	hiw^{5} taj^{1}	(mœŋ^{6})	tʃiaŋ^{1}
HJ	pɔŋ^{1}	hɛt^{9}	tɐj^{1}	tʰaj^{3}	tsɛŋ^{1}
FK	(tɔŋ^{2})	hik^{-33}	tɐj^{1}	(hɔn^{5})	(him^{5})
GZ	pɔŋ$^{1-53}$	(sik^{8})	!tsi$^{1-53}$	tʰɐj^{3}	tsaŋ$^{1-53}$
NH	pɔŋ^{1}	(sik^{8})	!tsi^{1}	tʰɐj^{3}	tsaŋ^{1}
DG	pɔŋ^{1}	(sik^{8})	tɐj^{1} tɔw^{5}	tʰɔj^{3}	tsɛŋ^{1}
FG	pɔŋ^{1}	(sɐk^{8})	(hiw^{3} tɐk^{7})	tʰɐj^{3}	tsaŋ^{1}
YF	pɔŋ^{1}	(sik^{8})	(tsi^{1} tuw^{2})	tʰɐj^{3}	tsaŋ^{1}
XH	pɔŋ^{1}	hæk^{9}	tij^{1}	hæj^{3}	tsaŋ^{1}
TS	pɔŋ^{1}	hiak^{9}	i^{1} tʰuj^{2}	haj^{3}	tsʰaŋ^{1}
KP	vɔŋ^{1}	hiak^{9}	ij^{1} tʰuj^{2}	haj^{3}	tsaŋ^{1}
EP	puaŋ^{1}	hiak^{9}	ti^{1}	haj^{3}	tsʰaŋ^{1}
BL	poŋ^{1}	hik^{7}	(ʃik^{7} tɐk^{7})	tʰɐj^{3}	tʃɛŋ^{1}
QZ	pɔŋ^{1}	(ʃik^{8})	(ʃik^{7} tɐk^{7})	tʰaj^{3}	tʃaŋ^{1}
NN_p	paŋ^{1}	hik^{9}	(ɬik^{9} tɐk^{9})	(han^{5})	tʃɛŋ^{1}
BS_p	paŋ^{1}	hɐt^{9}	(ʃət^{9} tɐk^{9})	(han^{5})	tʃɐŋ^{1}
BY	(tʃʰɐw^{5})	hɐt^{7}	(tʃʰuŋ^{1})	(hœn^{5})	tʃaŋ^{1}
MS	pɔŋ^{1}	hik^{7}	(ʃɐk^{7} liw^{4})	(hɔn^{5})	tʃɛŋ^{1}
YJ	pɔŋ^{1}	hɐt^{7}	!tʃi^{1} tɐk^{7}	(hɔn^{5})	tʃaŋ^{1}
WC	–	hɐk^{7}	–	tʰɐj^{3}	–

Table C.1

Gloss	pick up w/chopsticks (1)	pick up w/chopsticks (2)	search for	sleep	stroll
Index	6	7	8	9	10
Orth	挾	挾	揾	睏	蕩
CY	*kɛp^{7}	*[ŋ]ɛp^{7}	*un^{3}	*xun^{5}	*dɔŋ^{6}
YS	ket$^{9\text{-}55}$ ts^{h}oj$^{5\text{-}35}$	–	wɐn$^{\text{-}33}$	fɐn$^{\text{-}33}$	tɔŋ$^{\text{-}22}$
LX	kɛt^{8} ts^{h}œj^{5}	–	(san^{5})	fɐn^{5}	tœŋ^{6}
HJ	–	ŋɛt^{9} ts^{h}ɔj^{5}	wɐn^{3}	fɐn^{5}	tɔŋ^{6}
FK	–	nap^{2} ts^{h}ɔj^{5}	wɐn$^{\text{-}22}$	fin$^{\text{-}33}$	(lɐn^{3})
GZ	kap^{9} suŋ^{5}	–	wɐn^{3}	fɐn^{5}	tɔŋ^{6}
NH	kɛp^{9} suŋ^{5}	–	wɐn^{3}	hɐn^{5}	ʈɔŋ^{6}
DG	(k^{h}in^{2} suŋ^{5})	–	vɐn^{3}	fɐn^{5}	tɔŋ^{5}
FG	kap^{9} θuŋ^{5}	–	wɐn^{3}	fɐn^{1}	tɔŋ^{6}
YF	–	nap$^{9\text{-}22}$ ts^{h}u^{5}	wɐn^{2}	fɐn^{5}	(haŋ^{2})
XH	kap^{9} suŋ$^{1\text{-}21}$	–	væn^{3}	fæn^{1}	(law^{4})
TS	kap^{9} ɬuŋ$^{1\text{-}21}$	–	vun^{3}	fun^{1}	uŋ^{6}
KP	kap^{9} ɬuŋ$^{1\text{-}21}$	–	van^{3}	fun^{1}	uŋ^{6}
EP	kap^{9} suŋ$^{1\text{-}31}$	–	vun^{3}	fun^{1}	tuŋ^{6}
BL	kɛp^{9} t^{h}oj^{5}	–	wan^{3}	(ʃuj^{6})	tɔŋ^{6}
QZ	kɛp^{5} tʃ^{h}ɔj^{5}	–	wɐn^{3}	(ɬuj^{6})	(kɔŋ^{6})
NN_p	–	nɛp^{7} tʃ^{h}aj^{5}	(lɐj^{5})	(ɬuj^{6})	(jɐw^{2})
BS_p	–	nɛp^{7} tʃ^{h}aj^{5}	(lɐj^{5})	(ʃuj^{6})	(laŋ$^{\text{-}35}$)
BY	–	nɛp^{9} tʃ^{h}aj^{5}	(lœ^{3})	(ʃœ^{6})	(jəw^{2})
MS	kɛp^{55} t^{h}ɔj^{5}	–	(lɐn^{1})	(tʃuj^{6})	wuj^{5} tuaŋ$^{\text{-}24}$
YJ	–	–	wɐn^{3}	fɐn^{5}	(law^{5})
WC	kɛp^{8}	–	wɐn^{1}	(mɐj^{5})	t^{h}ɔŋ^{6}

Table C.2

Gloss	tell a story	bat	bear (1)	bear (2)	bird (1)
Index	11	12	13	14	15
Orth	講古	飛鼠	熊	熊	雀
CY	*kɔŋ³ ku³	*fi¹ ʃy³	*ɣuŋ²	*juŋ²	*tsiɔk⁷
YS	–	fɐj¹⁻⁴² sy³	hoŋ²	–	tsœk⁷ tsaj³
LX	–	fɐj⁻²² ʃy³	hɔŋ²	–	tsyk⁹ tsɐj³
HJ	–	fɐj¹ θy³	huŋ²	–	tsœk⁹ tsaj³
FK	–	fɐj⁻³³ sy³	–	jɔŋ²	tsɔk⁷ tsɐj³
GZ	kɔŋ³ ku³	fij¹⁻⁵⁵ sy³	huŋ²	–	tsœk⁹ tsɐj³
NH	kɔŋ³ ku³	pʰɐj¹ sy³	huŋ²	–	tœk⁹ ta³
DG	kɔŋ³ ku³ tsɔj³	fɐj¹ sy³	zuŋ²	zuŋ²	–
FG	–	fɐj¹ sy³	huŋ²	–	tsøk⁹ tsɐj³
YF	–	fij¹ sy³	–	juŋ²	tsœk⁹ tsɐj³
XH	kɔŋ³ ku³ tsæj³	fij¹ si³	huŋ²	–	tsiɔk⁷⁻²¹
TS	kɔŋ³ ku³ tɔj³	fij¹ si³	huŋ²⁻³⁵	–	tiak⁷⁻²¹
KP	kɔŋ³ ku³ tɔj³	fuj¹ si³	huŋ²⁻³⁵	–	tiak⁷⁻³⁵
EP	kuaŋ³ ku³ tsaj³	fuj¹ si³	huŋ²⁻³⁵	–	tsiɔk⁷⁻²¹
BL	kɔŋ³ ku³	fi¹ ʃy³	–	juŋ²	–
QZ	kɔŋ³ ku⁵ ɬʐ̩⁶	fi¹ ʃi³	–	juŋ²	–
NN_p	kaŋ³ kɔ¹	fuj¹ ɬuj³	–	juŋ²	–
BS_p	kaŋ³ kɔ³	fuj¹ ʃuj¹	–	jɔŋ²	–
BY	tʃaŋ³ kəw³	(pit¹ ʃuj³ ʃœŋ⁵)	–	juŋ²	–
MS	kuaŋ³ ku³	fi¹ ʃi³	–	juŋ²	–
YJ	kɔŋ³ ku⁵ ʃi⁶	fij¹ ʃi³	–	juŋ²	tʃiɛk⁹ tʃɐj³
WC	–	fi¹ ʃi³	ʃan¹ ɳɐn² huŋ²	ʃan¹ ɳɐn² huŋ²	tʃɔk⁹

Table C.3

Gloss	bird (2)	butterfly	cat	centipede (1)	chicken
Index	16	17	18	19	20
Orth	雀	蝴蝶	貓	百足	雞
CY	*tsʰiɔk^{7}	*wu^{2} diap^{8}	*ʔmɛw^{-55}	*pak^{7} tsuk^{7}	kɐj^{1}
YS	–	wu^{2} tet^{8}	mɐw^{-22}	(jy^{2} kuŋ^{1} tuŋ^{2})	kaj^{1}
LX	–	wu^{-22} tɛt^{8}	mɛw^{5}	(ɲi^{-22} kɔŋ^{3} tɔŋ^{2})	kaj^{1}
HJ	–	wu^{2} tit^{8}	miw^{1}	pɛk^{9} θuk^{7}	kaj^{1}
FK	–	wu^{2} tep^{8}	mɐw^{1}	(ŋ̍^{2} kɔŋ^{1})	kɐj^{1}
GZ	–	wu^{2} tip^{8}	maw^{1}	pak^{9} tsuk^{7}	kɐj^{1}
NH	–	fu^{2} tip^{8}	mɛw^{1}	pak^{9} tuk^{9}	kɐj^{1}
DG	tsʰø^{5} tsɔj^{3}	fu^{2} t[ik]^{8}	miw$^{1\text{-}55}$	pʰaw^{5} tsuk^{7}	kɔj^{1}
FG	–	wu^{2} tip^{8}	miw$^{1\text{-}35}$	pak^{9} tsuk^{7}	kɐj^{1}
YF	–	wu^{2} tit^{8}	miw^{1}	pak^{9} suk^{7}	kɐj^{1}
XH	–	vu^{2} tip^{8}	miw^{3}	pak^{9} tsuk^{7}	kæj^{1}
TS	–	vu^{2} iap^{8}	miw^{3}	!pak^{9} ɬuk^{7}	kaj^{1}
KP	–	vu^{2} ip^{8}	miw^{3}	(vak^{9} tʰuk^{7})	kaj^{1}
EP	–	–	miw^{3}	pak^{9} tsuk^{7}	kaj^{1}
BL	tʰɛk^{-33}	wu^{2} tɛp^{8}	m(i)ɛw^{1}	pik^{9} tuk^{7}	kɐj^{1}
QZ	tʃʰɛk^{9}	fu^{2} tip^{8}	mɛw^{1}	pak^{9} tʃuk^{7}	kɐj^{1}
NN_p	tʃʰɛk^{7}	hu^{2} tip^{8}	mɛw^{5}	pɛk^{9} tʃuk^{9}	kɐj^{1}
BS_p	tʃʰɛk^{7}	hu^{2} tip^{8}	mɛw^{5}	pɛk^{9} tʃɔk^{9}	kɐj^{1}
BY	ku^{4} tʃʰɛk^{9}	hu^{2} tip^{10}	mɛw^{1}	pak^{9} tʃuk^{7}	kɐj^{1}
MS	tʰiək^{7}	hu^{2} tip^{8}	mɛw^{1}	!pɛk^{55} ɬuk^{7}	kɐj^{1}
YJ	–	wu^{2} tip^{8}	miɛw^{5}	!pak^{9} ɬuk^{7}	kɐj^{1}
WC	–	–	mɛw^{1}	bak^{9} tuk^{7}	kɐj^{1}

Table C.4

Gloss	claws	cow	crab	crow	cuttlefish
Index	21	22	23	24	25
Orth	爪	牛	蟹	烏鴉	墨魚
CY	*tʃaw^{3}	*ŋɐw^{2}	*xaj[^{3}]	*u^{1} a^{1}	*mɐk^{8} ŋy^{2}
YS	tsew^{3}	ŋaw^{2}	ha^{3}	wu^{1} a^{1}	mɐk^{8} jy^{2}
LX	tʃɛw^{3}	ŋaw^{2}	haj^{4}	wu^{-22} a^{1}	mak^{8} ŋy^{2}
HJ	(ŋiw^{3})	ŋaw^{1}	(ha^{1} lat^{7})	wu^{1} a^{1}	mɐk^{8} jy^{2}
FK	(ɲaw^{3})	ŋɐw^{2}	(ha^{2} lat^{9})	(lɔ^{4} a^{1})	mɐk^{8} ɲy^{2}
GZ	tsaw^{3}	ŋɐw^{2}	haj^{-35}	wu^{1} a^{1}	mɐk^{8} jy^{2}
NH	tsaw^{3}	ŋɐw^{2}	!ha^{4}	wu^{1} a^{1}	mak^{8} jy^{2}
DG	tsaw^{3}	ŋaw^{2}	haj^{-35}	vu^{1} ŋa^{1}	mɐk^{8} zy$^{2-35}$
FG	tsaw^{3}	ŋɐw^{2}	(ha^{-33} lat^{8})	(liw^{3} a^{1})	mɐk^{8} jy^{2}
YF	tsaw^{3}	ŋɐw^{2}	(ha^{1} lak^{9})	wu^{1} a^{1}	mak^{8} jy^{2}
XH	tsaw^{3}	ŋaw^{2}	haj^{3}	vu^{1} a^{1}	mak^{8} ŋi$^{2-35}$
TS	tsaw^{3}	ŋew^{2}	haj^{3}	vu^{1} a$^{1-35}$	mak^{8} ŋuj$^{2-35}$
KP	tsaw^{3}	ŋaw^{2}	haj^{3}	vu^{1} a$^{1-35}$	mak^{8} ŋuj$^{2-35}$
EP	tsaw^{3}	ŋij^{2}	haj^{3}	vu^{1} a$^{1-35}$	mak^{8} ŋuj$^{2-35}$
BL	ɲaw^{3}	ɲɐw^{2}	!p^{h}oŋ^{2} haj^{4}	luk^{8} a^{1}	mɐk^{8} ɲy^{4}
QZ	tʃaw^{3}	ŋɐw^{2}	p^{h}ɔŋ^{2} haj^{3}	wu^{1} a^{1}	mɐk^{8} ɲi^{2}
NN_p	tʃaw^{3}	ɲɐw^{2}	(ha^{1} lat^{9})	ɔ^{1} ŋa^{1}	mɛk^{10} ɲuj^{2}
BS_p	tʃaw^{3}	ɲɐw^{2}	paŋ^{2} kaj^{3}	ɔ^{1} a^{1}	mɐk^{4} ɲuj^{2}
BY	(ɲaw^{3})	ɲ(i)əw^{2}	!paŋ^{2} haj^{5}	əw^{1} a^{1}	mɐk^{8} ɲuj^{2}
MS	tʃaw^{3}	ŋɐw^{2}	!puaŋ^{2} kaj^{4}	u^{1} a^{1}	mɐk^{8} ɲi^{2}
YJ	tʃaw^{3}	ŋɐw^{2}	haj^{3}	wu^{1} a^{1}	mɐk^{8} ji^{2}
WC	–	ɳɐw^{2}	–	–	–

Table C.5

Gloss	dog	duck	earthworm	egg (1)	egg (2)
Index	26	27	28	29	30
Orth	狗	鴨	黃蜎	蛋	舂
CY	*kɐw^{3}	*ap^{7}	*wɔŋ^{2} xiɔn^{3}	*dan^{6}	*tʃ^{h}yn
YS	kaw^{-44}	ap^{9}	hyn^{-35}	–	ts^{h}ɐn$^{1-52}$
LX	kaw^{6}	at^{9}	wɔŋ^{2} hyn^{4}	–	tʃ^{h}ɐn^{1}
HJ	kaw^{3}	ɛt^{9}	wɔŋ^{2} k^{h}un^{3}	–	ts^{h}ɐn^{1}
FK	kɔw^{3}	ap^{9}	(ɲɐn^{3})	–	ts^{h}ɐn^{1}
GZ	kɐw^{3}	ap^{9}	wɔŋ^{2} hyn^{3}	tan^{3}	ts^{h}œn^{1}
NH	kɐw^{3}	ɐp^{9}	ɭɔŋ^{2} hyn^{3}	–	ts^{h}œn^{1}
DG	kaw^{3}	ŋa^{5}	vɔ^{2} høn^{3}	–	ts^{h}ɐn^{1}
FG	kɐw^{3}	ap^{9}	wɔŋ^{2} hyn^{3}	tan^{-35}	ts^{h}ɐn^{1}
YF	kɐw^{3}	ap^{9}	wɔŋ^{2} hɐn^{3}	–	ts^{h}ɐn^{1}
XH	kæw^{3}	ap^{9}	vɔŋ^{2} hin^{3}	tan^{6}	–
TS	kew^{3}	ap^{9}	vɔŋ^{2} hun^{3}	an^{6}	–
KP	kaw^{3}	ap^{9}	vɔŋ^{2} hin^{3}	an^{6}	–
EP	kij^{3}	ap$^{9-35}$	vuaŋ^{2} hi[e]n^{3}	tan^{-35}	–
BL	kɐw^{3}	ap^{9}	(t^{h}u^{3} jɐn^{4})	tan^{6}	–
QZ	kɐw^{3}	ap^{9}	wɔŋ^{2} jyn^{3}	tan^{6}	–
NN_p	kɐw^{3}	ap^{9}	(t^{h}ɔ^{3} jɐn^{6})	tan^{6}	–
BS_p	kɐw^{3}	ap^{9}	(t^{h}ɔ^{3} jɐn^{6})	tan^{6}	–
BY	kɐw^{3}	ap^{9}	(t^{h}əw^{3} jən^{4})	tan^{6}	–
MS	kɐw^{3}	ap^{7}	(t^{h}u^{3} jən^{4})	tan^{6}	–
YJ	kɐw^{3}	ap^{9}	(k^{h}ɐw^{2} nɐn^{3})	tan^{6}	–
WC	kɐw^{3}	ap^{7}	(t^{h}u^{3} ŋoi^{6})	t^{h}an^{6}	–

Table C.6

Gloss	f. animal suff.	fish	fly	fox	frog
Index	31	32	33	34	35
Orth	𡊁 (女)	魚	蠅	狐狸	蛤
CY	*ʔna^{3}	*ŋy^{2}	*[ʔ]iŋ^{1}	*wu^{2} li^{2}	*kɐp^{7}
YS	na^{3}	jy^{2}	wu$^{1\text{-}52}$ jɐŋ^{2}	wu^{2} lɐj^{2}	kap$^{\text{-}33}$ na$^{\text{-}24}$
LX	na^{4}	ŋy^{2}	jiaŋ^{2}	wu$^{\text{-}22}$ lɐj^{2}	(tin^{2} kaj^{1})
HJ	na^{3}	jy^{2}	tsʰɔŋ^{1} jɐŋ^{2}	wu^{2} li^{3}	kɐp^{7} na^{3}
FK	(puə^{2})	ɲy^{2}	jɐŋ$^{\text{-}35}$	wu^{2} lɐj$^{2\text{-}35}$	kɐp^{7}
GZ	na^{3}	jy$^{2\text{-}35}$	wu$^{1\text{-}55}$ jiŋ$^{1\text{-}55}$	wu^{2} lij$^{2\text{-}35}$	(tʰin^{2} kɐj$^{1\text{-}55}$)
NH	na^{3}	jy^{2}	hiŋ^{2}	fu^{2} lɐj^{2}	(sœŋ^{1} hœŋ^{4})
DG	na^{3}	zy$^{2\text{-}35}$	ku^{1} ziŋ$^{\text{-}55}$	fu^{2} ŋɐj$^{2\text{-}35}$	kɔk^{7}
FG	na^{3}	jy^{2}	wu^{1} iŋ^{2}	wu^{2} lɐj^{2}	kap^{7} na^{3}
YF	na^{3}	jy^{2}	wu^{1} jiŋ^{1}	wu^{2} lij^{2}	kɐp^{7} na^{3}
XH	na^{3}	ŋi$^{2\text{-}35}$	vɔ^{2} ziŋ$^{\text{-}35}$	vu^{2} lij^{2}	kæp^{9}
TS	na^{3}	ŋuj$^{2\text{-}35}$	vɔ^{2} zen$^{\text{-}35}$	vu^{2} li$^{2\text{-}35}$	kep^{7} kaj^{3}
KP	na^{3}	ŋuj$^{2\text{-}35}$	vɔ^{2} ziŋ$^{\text{-}35}$	(mɐŋ^{3} kaw^{3})	(hin^{2} kaj$^{\text{-}35}$)
EP	na^{3}	ŋuj$^{2\text{-}35}$	vua^{2} zi(e)ŋ$^{\text{-}35}$	vu^{2} li$^{2\text{-}35}$	(hi(e)n^{2} kaj^{1})
BL	na^{3}	ɲy^{4}	mɐn^{1} jiŋ^{2}	wu^{2} li$^{\text{-}52}$	kɔp^{9} na^{3}
QZ	na^{3}	ɲi^{2}	mɐn^{1} jiŋ^{1}	fu^{2} li^{2}	(tʰin^{2} kɐj^{1})
NN_p	(mu^{4})	ɲuj^{2}	mɐn^{1} iŋ^{1}	hu^{2} luj^{2}	kap^{9}
BS_p	(mɐw^{4})	ɲuj^{2}	mɐn^{1} ən^{1}	(lɔŋ^{1} kɐw^{3})	kap^{9}
BY	(məw^{4})	ɲuj^{2}	mɐn^{2} jiŋ^{2}	la^{2} lɐj^{2}	kœp^{9}
MS	na^{3}	ɲi^{2}	wɐn^{2} iŋ^{1}	hu^{2} li^{2}	kəp^{7}
YJ	na^{3}	ji^{2}	(mɐn^{2} tʃij^{3})	wu^{2} lij^{2}	kɐp^{7}
WC	(nɛŋ^{1})	!ɲi^{2}	mɐm^{2} jiŋ^{2}	!fu^{2} li^{2}	kɔk^{9}

Table C.7

Gloss	goose	hare	hatch/brood chicks	hawk	hornet
Index	36	37	38	39	40
Orth	鵝	兔	孵雞仔	兀鷹	黃蜂
CY	*ŋɔ^{2}	*tʰu^{5}	*bɐw^{6} kɐj^{1} tsɔj^{3}	*ŋa^{2} iŋ^{1}	*wɔŋ^{2} fuŋ^{1}
YS	ŋɔ^{2}	tʰɐw$^{5-35}$	paw^{2} kaj^{1} tsaj^{3}	(taj^{6} jew^{4})	wɔŋ^{2} foŋ^{1}
LX	ŋɔ^{2}	tʰɐw^{5}	pɔ$^{-22}$ tsɐj^{3}	ŋa^{-22} jiaŋ^{1}	wɔŋ$^{-22}$ fɔŋ^{1}
HJ	ŋɔ^{2}	tʰɔw^{5}	pu^{6} kaj^{1} tsaj^{3}	(taj^{6} jiw^{6})	luŋ^{2} fuŋ^{1}
FK	ŋuə^{2}	tʰu^{5}	pɔ^{6} kɐj^{1} tsɐj^{3}	ŋa^{2} ɐŋ^{1}	(lɐŋ$^{-35}$)
GZ	ŋɔ$^{2-35}$	tʰuw^{5}	puw^{6} kɐj^{1} tsɐj^{3}	ma^{2} jiŋ^{1}	wɔŋ^{2} fuŋ^{1}
NH	ŋɔ^{2}	tʰuw^{5}	puw^{6} kɐj^{1} ta^{3}	ma^{2} jiŋ^{1}	pʰuŋ^{1} ʈɛŋ^{1}
DG	ŋɔ$^{2-35}$	tʰɔw^{5}	puw^{5} kɔj^{1} tsɐj^{3}	ŋa^{2} ziŋ^{1}	vɔŋ^{2} fuŋ$^{1-55}$
FG	ŋɔ^{2}	tʰu^{1}	paw^{6} kɐj^{1} tsɐj^{3}	ma^{2} jiŋ^{1}	wɔŋ^{2} fuŋ^{1}
YF	ŋɔ^{2}	tʰuw^{5} tsɐj^{3}	puw^{2} kɐj^{1} tsɐj^{3}	taj^{2} ŋa^{2} jiŋ^{1}	wɔŋ^{2} fuŋ^{1}
XH	ŋɔ$^{2-35}$	hæw^{1}	puw^{6} kæj^{1} tsæj^{3}	ŋa^{2} ziŋ^{1}	vɔŋ^{2} fuŋ^{1}
TS	ŋɔ$^{2-35}$	hu^{2}	pu^{6} kaj^{1} tɔj^{3}	ŋew^{2} zen^{1}	vɔŋ^{2} fuŋ$^{1-35}$
KP	ŋu$^{2-35}$	hu^{1}	vu^{6} kaj^{1} tɔj^{3}	ŋaw^{2} ziŋ$^{1-35}$	vɔŋ^{2} fuŋ$^{1-35}$
EP	ŋua$^{2-35}$	hu^{1}	pij^{6} kaj^{1} tsaj^{3}	ŋa^{2} ziŋ^{1}	vuaŋ^{2} fuŋ$^{1-35}$
BL	ŋɔ^{2}	tʰu^{5}	pɐw^{6} kɐj^{1} ɲi^{1}	ŋaj^{2} iŋ^{1}	ɔŋ^{2} fuŋ^{1}
QZ	ŋɔ^{3}	tʰu^{5}	pɐw^{6} kɐj^{1}	ŋa^{2} jiŋ^{1}	wɔŋ^{2} fuŋ^{1}
NN_p	ŋɔ^{2}	tʰu^{5}	paw^{6} kɐj^{1} ɲi^{2}	ŋaj^{2} iŋ^{1}	huŋ^{2} fuŋ^{1}
BS_p	ŋa^{2}	tʰɔ^{5}	paw^{6} kɐj^{1} ɲi^{2}	(mɛw^{2} ən^{1})	huŋ^{2} fɔŋ^{1}
BY	ŋœ^{2}	tʰəw^{5}	pœ^{6} kɐj^{1}	ŋa^{2} jiŋ^{1}	uŋ^{2} fuŋ^{1}
MS	ŋɔ^{2}	tʰu^{5}	pəw^{6} kɐj^{1} ŋɛ^{1}	ŋaj^{2} iŋ^{1}	hɔŋ^{2} fuŋ^{1}
YJ	ŋɔ^{2}	tʰɔw^{5} tʃɐj^{3}	pɔw^{6}	lɔw^{3} iŋ^{1}	wɔŋ^{2} fuŋ^{1}
WC	tʰin^{1} ŋɔ^{1}	–	pʰɔw^{6}	ŋak^{8} iŋ^{1}	!fuŋ^{1}

Table C.8

Gloss	horse	louse	m. animal suff.	monkey	mosquito
Index	41	42	43	44	45
Orth	馬	蝨	公	馬騮	蚊
CY	ma^{4b}	*ʃit^{7}	*kuŋ^{1}	*ʔma^{-55} ʔlɐw^{-55}	*ʔmun^{1}
YS	ma^{4}	sɐt^{7} na^{3}	(ku^{3})	ma^{4} lɐw^{1}	mɐn^{1} tsaj^{3}
LX	ma^{-22}	kaw^{5} ʃɐt^{7}	(ku^{3})	ma^{-22} lɐw^{1}	mɐn^{-22} tsɐj^{3}
HJ	ma^{4}	θɐt^{7} na^{3}	kuŋ^{1}	ma^{4} law^{1}	mɐn^{1} tsaj^{3}
FK	ma^{4}	sɐt^{7}	(ku^{3})	ma^{1} lɐw^{1}	mɐn^{1} tsɐj^{3}
GZ	ma^{4}	sɐt^{7} na^{-35}	kuŋ$^{1-55}$	ma^{4} lɐw^{1}	mɐn^{1}
NH	ma^{4}	sɐt^{7} na^{-35}	kuŋ^{1}	ma^{4} lɐw^{1}	mɐn^{2}
DG	ma^{4}	sɐk^{7} na^{-35}	kuŋ^{1}	ma^{4} ŋɐw$^{1-55}$	mɐn$^{1-55}$
FG	ma^{5}	sɐt^{7} na^{-35}	(ku^{3})	ma^{1} lɐw^{1}	mɐn^{1} tsɐj^{3}
YF	ma^{4}	sɐt^{7} na^{-35}	(ku^{3})	ma^{4} lɐw^{1}	mɐn^{1} /mɐn^{2} tsɐj^{3}
XH	ma^{4}	sæt^{7} na^{3}	(ku^{3})	ma^{3} læw^{3}	mæn^{-35}
TS	ma^{4}		kuŋ$^{-21}$	ma^{3} lew^{3}	mun^{-35}
KP	ma^{4}	sit^{7} na^{3}	kuŋ$^{-21}$	ma^{3} lew^{3}	mun^{-35}
EP	ma^{6}	si[e]t^{7} na^{3}	kuŋ$^{-31}$	ma^{3} lij^{3}	mian^{-35}
BL	ma^{4}	ʃɐt^{7}	kuŋ^{1}	ma^{1} lɐw^{1}	mɐn^{1} ti^{3}
QZ	ma^{4}	ʃɐt^{7}	kuŋ^{1}	ma^{1} lɐw^{1}	mɐn^{1} tʃi^{3}
NN_p	ma^{4}	ɬɐt^{9}	kuŋ^{1}	ma^{1} lɐw^{1}	mɐn^{1} tʃi^{3}
BS_p	ma^{4}	ɬɐt^{9}	kɔŋ^{1}	ma^{1} lɐw^{1}	mɐn^{1} tʃi^{3}
BY	ma^{4}	ɬɐt^{7}	kuŋ^{1}	ma^{1} ləw^{1}	mɐn^{2} tʃəj^{3}
MS	ma^{4}	ʃət^{7}	(ku^{3})	ma^{1} lɐw^{1}	wɐn^{2} ɬi^{3}
YJ	ma^{3}	ʃɐt^{7} na^{3}	kuŋ^{1}	ma^{3} lɐw^{2} tʃɐj^{3}	mɐn^{2} tʃ^{h}uŋ^{2}
WC	–	ʃɐt^{7}	(ku^{1})	ma^{4} lɐw^{1}	mɐn^{1}

Table C.9

Gloss	mud-eel (1)	mud-eel (2)	myna bird	owl	parrot
Index	46	47	48	49	50
Orth	鱔	鱔	鶇哥	貓頭鷹	鸚鵡
CY	*ʒian[4a]	*ʒian^{6}	*ʔlɛw^{-55} kɔ^{1}	*ʔmɛw^{-55} dɐw^{2} iŋ^{1}	*iŋ^{1} mu^{4b}
YS	wɔŋ^{2} sin^{4}	–	lew^{-33} kɔ^{1}	mew^{5} taw^{2} jɐŋ^{1}	jɐŋ^{1} mu^{3}
LX	ʃin^{4}	–	lɛw^{-22} kɔ^{1}	mɛw^{-22} taw^{-22} jiaŋ^{1}	jiŋ^{1} mu^{4}
HJ	–	θin^{6}	ha^{6} liw^{1}	(kʰɐm^{3} muk^{8} tsœk^{9})	jɛŋ^{1} mu^{4}
FK	uŋ^{2} tsin^{-35}	–	lew^{1} kœ^{1}	mew^{1} tɐw^{2} ɐŋ^{1}	jiŋ^{1} mu^{4}
GZ	wɔŋ^{2} sin^{-35}	–	liw^{1} kɔ^{1}	maw^{1} tʰɐw^{2} jiŋ^{1}	jiŋ^{1} kɔ^{1}
NH	fɔŋ^{2} sin^{4}	–	tiw^{5} kɔ^{1}	mɛw^{1} tʰaw^{2} jiŋ^{1}	jiŋ^{1} muw^{4}
DG	sin^{4}	–	liw$^{1-55}$ kɔ$^{1-55}$	miw^{1} tʰaw^{2} ziŋ^{1}	zɔŋ^{1} mɔw^{4}
FG	wɔŋ^{2} sin^{4}	–	liw^{1} kɔ^{1}	miw^{1} tʰɐw^{2} jiŋ^{1}	jiŋ^{1} mu^{3}
YF	–	wɔŋ^{2} sin^{6}	liw^{4} kɔ^{1}	miw^{1} tʰɐw^{2} jiŋ^{1}	jiŋ^{1} mu^{4}
XH	–	sin^{-21}	liw^{4} kɔ^{1}	miw^{3} hæw^{2} ziŋ^{1}	aŋ^{1} mæw^{4}
TS	–	sian^{6}	liaw^{4} kɔ^{1}	miw^{3} hew^{2} zen$^{1-35}$	en^{1} mu^{-35}
KP	–	siŋ^{6}	liw^{3} ku^{3}	miw^{3} haw^{2} ziŋ$^{1-35}$	ziŋ^{1} mu^{-35}
EP	–	sian^{6}	liw^{3} kua^{3}	miw^{3} hij^{2} ziŋ$^{1-35}$	ziŋ^{1} mu^{6}
BL	–	ɔŋ^{2} ʃin^{6} ɲy^{4}	lɛw^{1} kɔ^{1}	m(i)ɛw^{1} tʰɐw^{2} tʰɛk^{-33}	(lɛw^{4} kɔ^{1})
QZ	wɔŋ^{2} ʃin^{3} ɲi^{2}	–	lɛw^{1} kɔ^{1}	mɛw^{1} tʰɐw^{2} jiŋ^{1}	jiŋ^{1} mu^{4}
NN_p	huŋ^{2} ɬin^{-243}ɲuj^{2}	–	lɛw^{1} kɔ^{1}	mɛw^{5} tɐw^{2} iŋ^{1}	iŋ^{1} mu^{4}
BS_p	–	huŋ^{2} ʃin^{6} ɲuj^{2}	(paw^{-35} pa^{1})	mɛw^{5} tɐw^{2} iŋ^{1}	(ən^{1} fu^{4})
BY	–	uŋ^{2} ʃin^{6} ɲuj^{2}	liw^{3} kœ^{1}	mɛw^{2} tɐw^{2} jiŋ^{1}	(jiŋ^{1} fəw^{4})
MS	hɔŋ^{2} ɬin^{4} ɲi^{2}	–	liw^{1} kɔ^{1}	(əm^{5} ʃan^{1} ŋaj^{2})	(iŋ^{1} hu^{4})
YJ	wɔŋ^{2} ʃin^{3}	–	liɛw^{5} kɔ^{5}	miɛw^{5} tʰɐw^{2} iŋ^{1}	iŋ^{1} mɔw^{3}
WC	–	–	liw^{1} kɔ^{1}	–	iŋ^{1} kɔ^{1}

Table C.10

Gloss	pig	pigeon	rat	ribbonfish	scale (fish) (1)
Index	51	52	53	54	55
Orth	豬	白鴿	老鼠	帶	魚鱗
CY	*tʃy^{1}	*bak^{8} kɔp^{7}	*lɔw^{4b} ʃy^{3}	*taj^{5}	*ŋy^{2} lyn^{2}
YS	!tœj^{1}	pek^{8} kap^{7}	lo$^{4-33}$ sy^{3}	–	jy^{2} lɐn^{2}
LX	!tœj^{-22}	piak^{8} kat^{7}	lɔw^{-22} ʃy^{3}	taj^{5} ŋy^{2}	ŋy^{-22} lɐn^{2}
HJ	!tœj^{1}	pɛk^{8} kap^{9}	lɔw^{4} θy^{3}	taj^{5} jy^{2}	jy^{2} lɐn^{2}
FK	!tœj^{1}	pek^{8} kɔp^{9}	lɔ^{4} sy^{3}	taj^{5} ɲy$^{2-35}$	ɲy^{2} lɐn^{-35}
GZ	tsy^{1}	pak^{8} kɐp$^{9-35}$	luw^{4} sy^{3}	ŋa^{2} taj$^{5-35}$	jy^{2} lœn^{2}
NH	tsy^{1}	pak^{8} kup^{9}	lɔ^{4} sy^{3}	jiw^{1} ta^{5} jy^{2}	jy^{2} lœn^{2}
DG	tsy^{1}	pɛk^{8} ka^{5}	ŋɔw^{4} sy^{3}	ŋa^{2} taj^{5}	zy^{2} ŋɐn^{2}
FG	tsy^{1}	pak^{8} kap^{7}	law$^{3-35}$ sy^{3}	taj^{1} jy^{2}	jy^{2} lɐn^{2}
YF	tsy^{1}	pak^{8} kɐp^{7}	luw^{4} sy^{3}	ts^{h}ɛ^{1} taj^{5} jy^{2}	–
XH	tsi^{1}	pak^{8} kap$^{9-11}$	lu^{1} si^{3}	taj^{1} ŋi$^{2-35}$	–
TS	tsi^{1}	pak^{8} ap^{-11}	law^{1} si^{3}	ziaw^{1} aj$^{1-35}$	–
KP	tsi^{1}	vak^{8} ap^{-21}	lɔ^{1} si^{3}	ziaw^{1} aj$^{1-35}$	–
EP	tsi^{1}	pak^{8} ap$^{9-35}$	law^{1} si^{3}	pak^{8} taj^{1}	–
BL	tʃy^{1}	pɛk^{-22} kɐp^{9}	lɐw^{4} ʃy^{3}	taj^{5} ɲy^{4}	ɲy^{4} lɐn^{2}
QZ	tʃi^{1}	pak^{8} kap^{9}	lɐw^{4} ʃi^{3}	taj^{5} ɲi^{2}	ɲi^{2} lɐn^{2}
NN_p	tʃuj^{1}	puk^{8} kap^{9}	law^{4} ɬuj^{3}	taj^{5} ɲuj^{2}	ɲuj^{2} lɐn^{2}
BS_p	tʃuj^{1}	puk^{8} kap^{9}	law^{4} ʃuj^{1}	taj^{5} ɲuj^{2}	ɲuj^{2} lɐn^{2}
BY	tʃuj^{1}	pak^{8} kœp^{9}	lœ^{4} ʃuj^{3}	taj^{5} ɲuj^{2}	ɲuj^{2} lən^{2}
MS	tʃi^{1}	kəp^{7}	ləw^{4} ʃi^{3}	taj^{5} ɲi^{2}	ɲi^{2} lɐn^{2}
YJ	tʃi^{1}	pak^{8} kap^{9}	lɔw^{3} ʃi^{3}	taj^{5} ji^{2}	ji^{2} lɐn^{2}
WC	tʃi^{3}	–	ʃi^{3}	–	–

Table C.11

Gloss	scale (fish) (2)	shrimp	silkworm	snake	tail
Index	56	57	58	59	60
Orth	魚鱗	蝦	蠶	蛇	尾
CY	*ŋy^{2} ʔlin^{1}	*xa^{1}	dzam^{2}	ʒia^{2}	*ʔmi^{3}
YS	–	ha$^{1\text{-}52}$	tsam^{2} toŋ^{2}	se^{2}	mɐj^{3}
LX	–	ha^{1}	tsan^{2}	ʃɛ^{2}	mɐj^{3}
HJ	–	ha^{1} kuŋ^{1}	tsam^{2}	tsɛ^{2}	mɐj^{3}
FK	–	ha^{1}	tsam^{2}	tsiə^{2}	mɐj^{4}
GZ	–	ha$^{1\text{-}53}$	ts^{h}am^{2}	sɛ^{2}	mij^{4}
NH	–	ha^{1}	tam^{2}	tsɛ^{2}	mɐj^{4}
DG	–	ha^{1} kuŋ$^{\text{-}55}$	ts^{h}aŋ^{2} ts^{h}uŋ$^{2\text{-}35}$	sø^{2}	mɐj^{4}
FG	–	ha^{1}	ts^{h}am^{2}	se^{2}	mɐj^{6}
YF	jy^{2} lɐn^{1}	ha^{5} kuŋ^{1}	ts^{h}am^{2}	sɛ^{2}	mij^{4}
XH	ŋi^{2} læn^{1}	ha^{1}	ts^{h}am$^{\text{-}21}$	sæ^{2}	mij^{3}
TS	ŋuj^{2} lin^{1}	ha$^{\text{-}35}$	t^{h}am$^{\text{-}35}$	sia^{2}	m[e]i^{3}
KP	ŋuj^{2} liŋ^{1}	ha$^{\text{-}35}$	t^{h}am$^{2\text{-}35}$	sia^{2}	mij^{3}
EP	ŋuj^{2} lian^{1}	ha$^{\text{-}35}$	ts^{h}am$^{2\text{-}35}$	sia^{2}	mi^{3}
BL	–	ha^{1} kuŋ^{1}	t^{h}am^{2}	ʃɛ^{2}	mi^{4}
QZ	–	ha^{1}	tʃ^{h}am^{2}	ʃɛ^{2}	mi^{4}
NN_p	–	ha^{1} tʃi^{3}	tʃan^{2}	ɬi^{2}	muj^{4}
BS_p	–	ha^{1}	tʃan^{2}	ʃi^{2}	muj^{4}
BY	–	ha^{1} kuŋ^{1}	tʃan^{2}	ʃi^{2}	məj^{4}
MS	–	ha^{1}	ɬam^{2}	tʃɛ^{2}	mi^{4}
YJ	–	ha^{1}	tʃ^{h}am^{2}	ʃiɛ^{2}	mij^{3}
WC	–	ha^{1} kuŋ^{1}	–	ʃɛ^{2}	mi^{4}

Table C.12

Gloss	tiger	toad	wild goose	wolf	dad
Index	61	62	63	64	65
Orth	老虎	蟾蜍	雁	狼	爸
CY	*lɔw^{4a} xu^{3}	*gɐm^{2} dʒy^{2}	*ŋan^{6}	*lɔŋ^{2}	*pa^{1}
YS	lo$^{4\text{-}33}$ fu^{3}	kɐm$^{\text{-}33}$ tsy^{2}	ŋen^{6}	lɔŋ^{2}	a$^{\text{-}44}$ pa$^{1\text{-}52}$
LX	lɔw$^{\text{-}22}$ fu^{3}	!kɐn^{6} kœj^{2}	ŋɛn^{6}	lœŋ^{2}	a$^{\text{-}22}$ pa^{1}
HJ	lɔw^{4} fu^{3}	k^{h}ɐm^{1} tsy^{2}	ŋɛn^{6}	lɔŋ^{2}	a$^{\text{-}33}$ pa^{1}
FK	lɔ^{4} fu^{3}	!kam$^{\text{-}22}$ ky^{2}	ŋan^{6}	lœŋ^{2}	ŋ̍$^{\text{-}22}$ pa^{3}
GZ	luw^{4} fu^{3}	!k^{h}ɐm^{2} k^{h}œj^{3}	ŋan^{6}	lɔŋ^{2}	a$^{\text{-}33}$ pa$^{1\text{-}55}$
NH	lɔ^{4} fu^{3}	k^{h}ɐm$^{\text{-}21}$ tsy^{2}	ŋaŋ^{6}	lɔŋ^{2}	–
DG	ŋɔw^{4} fu^{3}	k^{h}ɐm^{2} sy^{3}	ŋɛŋ^{5}	ŋɔŋ^{2}	a^{5} pa^{2}
FG	law$^{3\text{-}33}$ fu^{3}	k^{h}ɐm^{2} k^{h}y^{2}	ŋan^{6} ŋɔ^{2}	lɔŋ^{2}	a^{1} pa^{1}
YF	luw^{4} fu^{3}	!k^{h}ɐm^{2} k^{h}œj^{2}	(ŋa^{2} jiŋ$^{1\text{-}35}$)	lɔŋ^{2}	pa^{5} pa^{1}
XH	luw^{1} fu^{3}	k^{h}æm^{2} si^{2}	ŋan^{6}	lɔŋ^{2}	a^{1} pa^{3}
TS	law^{1} fu^{3}	k^{h}em^{2} si^{2}	ŋan^{6}	lɔŋ$^{2\text{-}35}$	a^{1} pa^{3}
KP	lɔ^{1} fu^{3}	k^{h}am^{2} si^{2}	ŋan^{6}	lɔŋ$^{2\text{-}35}$	a^{1} ia^{1}
EP	law^{1} fu^{3}	k^{h}iam^{2} si^{2}	ŋan$^{6\text{-}35}$	luaŋ$^{2\text{-}35}$	a^{1} tia$^{\text{-}35}$
BL	(taj^{6} tʃ^{h}uŋ^{2})	k^{h}ɐm^{2} ʃy$^{\text{-}22}$	ŋan^{6}	lɔŋ^{2}	pa^{1}
QZ	lɐw^{4} fu^{3}	k^{h}ɐm^{2} ʃi^{2}	ŋan^{6}	lɔŋ^{2}	pa^{1}
NN_p	law^{4} hɔ^{3}	ɬɐm^{2} ɬuj^{2}	ɲan^{6}	laŋ^{2}	pa^{1}
BS_p	law^{4} hɔ^{3}	ʃɐm^{2} ʃuj^{2}	ɲan^{6}	laŋ^{2}	pa^{1}
BY	(taj^{6} tʃuŋ^{2})	!k^{h}ɐm^{2} k^{h}e^{2}	ɲan^{6}	lœŋ^{2}	ji^{1}
MS	ləw^{4} hu^{3}	k^{h}əm^{2} tʃi^{2}	ŋan^{6}	luaŋ^{2}	–
YJ	lɔw^{3} fu^{3}	kɐm^{5} ʃi^{2}	taj^{6} ŋɔn^{6}	lɔŋ^{2}	a^{1} pa^{5}
WC	(t^{h}aj^{6} mɛw^{1})	k^{h}ɐm^{2} ʃi^{2}	(t^{h}in^{1} ŋɔ^{1})	–	–

Table C.13

Gloss	daughter	daughter-in-law (1)	daughter-in-law (2)	grandfather	grandmother
Index	66	67	68	69	70
Orth	女	新婦	新婦	公	婆
CY	*ʔny^{3}	*sin^{1} bu[^{4a}]	*sin^{1} fu^{4b}	*kuŋ^{1}	*bɔ^{2}
YS	nɔj^{-24}	sɐn$^{1-42}$ pu^{-24}	–	a$^{1-42}$ koŋ$^{1-52}$	a$^{1-42}$ pɔ^{2}
LX	nœj^{4}	sɐn^{-22} pɔw^{4}	–	a^{-22} kɔŋ^{1}	a^{-22} pɔ^{2}
HJ	nœj^{-35}	θɐn^{1} pu^{6}	–	a^{5} kuŋ^{1}	a^{5} pɔ^{2}
FK	nœj^{4}	sɐn^{-33} pu^{4}	–	ɔŋ^{1}	puə^{2}
GZ	nœj^{3}	sɐm$^{1-53}$ pʰuw^{4}	–	a^{5} kuŋ$^{1-55}$	ŋɔj^{6} pʰɔ^{2}
NH	nœj^{3}	sɐn^{1} muw^{4}	–	a^{5} kuŋ^{1}	a^{5} pʰɔ^{2}
DG	nuj^{3}	sɐŋ^{1} pʰɔw^{4}	–	a^{5} kuŋ^{1}	a^{5} pʰɔ^{2}
FG	nøj^{3}	θɐn^{1} pʰu^{4}	–	a^{1} kuŋ^{1}	taj^{1} pʰɔ^{3}
YF	nœj^{4}	sɐn^{1} pʰu^{4}	–	a^{5} kuŋ^{1}	a^{5} pʰɔ^{2}
XH	nuj^{3}	–	sæn^{1} fu^{4}	a^{1} kuŋ^{1}	a^{1} pʰɔ^{2}
TS	nuj^{3}	–	ɬin^{1} fu^{4}	a^{1} kuŋ^{1}	a^{1} pʰɔ^{2}
KP	nuj^{3}	–	ɬiŋ^{1} fu^{4}	a^{1} kuŋ^{1}	a^{1} hu^{2}
EP	nuj^{3}	–	si(e)n^{1} fu^{6}	kuŋ^{1} taj^{-35}	pʰua^{2} taj^{-35}
BL	ny^{4}	–	ɬɐn^{1} fu^{4}	kuŋ^{1}	pʰɔ^{2}
QZ	ny^{3}	ɬɐn^{1} pʰɐw^{4}	–	kuŋ^{1}	pʰɔ^{2}
NN_p	nuj^{4}	–	fɐn^{1} fɐw^{-24}	kuŋ^{1}	pu^{2}
BS_p	nuj^{4}	–	fɐn^{1} fɐw^{4}	kɔŋ^{1}	pu^{2}
BY	nuj^{4}	–	ɬən^{1} fəw^{4}	kuŋ^{1}	pəw^{2}
MS	nuj^{-24}	ɬɐn^{1} pəw^{4}	–	kuŋ^{1}	ma^{1}
YJ	fun^{2} tʃɐj^{3}	–	ɬɐn^{1} fu^{3}	–	ka^{1} pʰɔ^{2}
WC	ni^{2}	–	ɬɐn^{1} fu^{4}	a^{5} kuŋ^{1}	a^{5} pʰɔ^{2}

Table C.14

Gloss	grandson	in-law/maternal	little sister	maternal uncle	mom
Index	71	72	73	74	75
Orth	孫	外	妹	舅	媽
CY	*sun^{1}	*ŋ^{w}ɔj^{6}	*muɔj^{6}	*giw^{4b}	*ma^{-55}
YS	syn$^{1-52}$	ŋoj^{-22} mu^{-24}	a$^{1-42}$ muj^{-22}	a$^{1-42}$ kɐw^{-24}	a$^{1-42}$ am^{3}
LX	syn^{3}	–	a^{-22} muj^{-22}	kɐw^{4}	a^{-22} ma^{3}
HJ	θyn^{1}	ŋɔj^{6} mu^{4}	a^{5} muj^{6}	tsaw^{6} fu^{6}	a^{5} ma^{1}
FK	ŋ̍$^{-22}$ sɐn^{1}	–	siə^{3} muj^{6}	tsɔw^{-22} ɔŋ^{1}	a^{-22} ma^{3}
GZ	syn$^{1-55}$	ŋɔj^{6} muw^{3}	a^{5} muj^{-35}	kʰɐw^{4} fu^{3}	a^{-33} ma$^{1-55}$
NH	syn^{1}	ŋɔ^{6} muw^{4} n^{3}	a^{5} my^{-35}	kʰɐw^{4} fu^{6}	a^{-33} ma^{1}
DG	søn^{-55}	–	muj^{5}	kaw^{5} fu^{3}	a^{5} ma^{-55}
FG	θyn^{1}	ŋøj^{6} mu^{4}	a^{1} muj^{-35}	kʰɐw^{2} fu^{3}	a^{1} ma^{1}
YF	syn^{1}	ŋu^{2} mu^{4} na^{3}	a^{5} mu^{-35}	kʰɐw^{2} fu^{3}	a^{5} ma^{1}
XH	sæn^{1}	ŋuj^{6} mu^{-35}	a^{1} muj^{-21}	a^{1} kʰæw^{4}	a^{1} ma^{3}
TS	ɬun^{1}	ŋɔj^{6} mu^{-35}	!a^{1} mɔj^{3}	a^{1} kʰiw^{4}	a^{1} ma^{3}
KP	ɬun^{1}	ŋɔj^{6} mu^{-35}	a^{1} mɔj^{-35}	a^{1} kʰiw^{4}	a^{1} ma^{3}
EP	sun^{1}	ŋuaj^{6} mu^{-35}	a^{1} muaj^{-35}	a^{1} kʰij^{-35}	a^{1} ma^{3}
BL	ɬyn^{1}	ŋaj^{6} mu^{4}	muj^{6}	kʰɐw^{4} fu^{6}	ma^{1}
QZ	ɬin^{1}	ŋuj^{6} mu^{4}	muj^{3}	kʰɐw^{4} fu^{6}	ma^{1}
NN_p	ɬɔn^{1}	waj^{6} pu^{2}	(kɔ^{5} nɛŋ^{2})	kɐw^{4} jɛ^{2}	ma^{1}
BS_p	ɬɔn^{1}	waj^{6} mɐw^{4}	muj^{6}	kɐw^{4} jɛ^{2}	ma^{1}
BY	ɬun^{1}	–	(kœ^{6} nɛŋ^{6})	tʃəw^{4} kuŋ^{1}	na^{3}
MS	ɬun^{1}	ŋuj^{6} ləw^{3}	ləw^{4} muj^{6}	tʃɐw^{4} fu^{4}	(əw^{1})
YJ	ɬun^{1}	wuj^{6} mɔw^{3}	ɬiw^{1} muj^{6}	a^{1} kʰiɛw^{3}	a^{1} ma^{5}
WC	ɬun^{1}	–	muj^{6}	a^{5} kʰɐw^{4}	a^{5} ma^{1}

Table C.15

Gloss	older paternal uncle	son	wife of younger paternal uncle	younger paternal uncle	alley
Index	76	77	78	79	80
Orth	伯	仔	嬸	叔	巷
CY	*pak^{7}	*ts[ɐ]j^{3}	*ʃim^{3}	*ʃuk^{7}	*ɣɔŋ^{6}
YS	a$^{1\text{-}42}$ pek^{7}	tsaj^{3}	a$^{1\text{-}42}$ sɐm^{3}	a$^{1\text{-}42}$ sok^{7}	hɔŋ$^{-22}$
LX	a^{-22} piak^{7}	tsaj^{3}	a^{-22} ʃɐm^{3}	a^{-22} ʃɔk^{7}	hœŋ$^{-22}$
HJ	(tɛ^{1})	tsaj^{3}	θɐm^{3}	a^{5} θuk^{7}	kaj^{1} hɔŋ^{6}
FK	taj^{6} pa^{3}	tsɐj^{3}	sɐm^{3}	sɔk^{7}	œŋ$^{-35}$
GZ	a^{-33} pak^{9}	tsɐj^{3}	a^{5} sɐm^{3}	a^{-33} suk^{7}	hɔŋ$^{-35}$
NH	a^{-33} pak^{9}	ta^{3}	a^{5} sɐm^{3}	a^{-33} suk^{7}	hɔŋ$^{-35}$
DG	pɛ$^{-224}$ zø^{2}	tsɔj^{3}	a^{5} sɐm^{3}	a^{5} suk^{7}	hɔŋ^{5}
FG	a^{1} pak^{9}	tsɐj^{3}	a^{1} sɐm^{3}	a^{1} suk^{7}	hɔŋ$^{-35}$
YF	a^{5} pak^{9}	tsɐj^{3}	a^{5} sɐm^{3}	a^{5} suk^{7}	hɔŋ^{2} tsɐj^{3}
XH	a^{1} pak^{9}	tsæj^{3}	a^{1} sæm^{3}	a^{1} suk^{7}	hɔŋ^{6}
TS	a^{1} pak^{-11}	tɔj^{3}	a^{1} sim^{3}	a^{1} suk^{3}	hɔŋ^{6}
KP	a^{1} vak^{8}	tɔj^{3}	a^{1} sim^{3}	a^{1} suk^{7}	hɔŋ$^{-35}$
EP	pak^{7} tia^{-35}	tsaj^{3}	a^{1} si(ə)m^{3}	a^{1} suk^{7}	huaŋ^{6}
BL	pak^{9}	tʃɐj^{3}	ʃɐm^{3}	ʃuk^{7}	hɔŋ^{6}
QZ	pak^{9}	tʃɐj^{3}	ʃɐm^{3}	ɬuk^{7}	hɔŋ^{6}
NN_p	(jɛ^{2})	tʃaj^{3}	jɐm^{3}	ɬuk^{9}	haŋ^{6} tʃi^{3}
BS_p	pɛk^{9} jɛ^{2}	(ɲi^{2})	ɬɐm^{3}	ʃɔk^{9}	haŋ^{6} tʃi^{3}
BY	(taj^{6} jɛ^{2})	tʃaj^{3}	ʃɐm^{3}	ʃuk^{7}	haŋ^{6}
MS	(tɛ^{1})	(ŋi^{2})	ʃɐm^{3}	ʃuk^{7}	wuaŋ^{6}
YJ	a^{1} pak^{9}	tʃɐj^{3}	–	a^{1} ʃuk^{7}	hɔŋ^{6}
WC	bak^{9} dɛ^{1}	tɐj^{3}	a^{5} ʃɐm^{3}	a^{5} ʃuk^{7}	haŋ^{6} ɲi^{2}

Table C.16

Gloss	bottle	chopping board (1)	chopping board (2)	chopsticks	corner
Index	81	82	83	84	85
Orth	樽	砧板	砧板	筷	角落
CY	*tsyn^{1}	*tʃim^{1} pan^{3}	*tim^{1} pan^{3}	*k^{wh}aj^{5}	*kɔk^{7} lɔk^{7}
YS	tsɐn$^{1\text{-}52}$	–	tɐm^{1} pan^{3}	(tœj$^{\text{-}22}$)	kɔk^{3} lɔk$^{\text{-}22}$ taw^{2}
LX	tsɐn^{1}	–	tɐm$^{\text{-}22}$ pan^{3}	k^{wh}aj^{5} tsɐj^{3}	kœk^{9} lyk^{7}
HJ	tsɐn^{1}	–	tɐm^{1} pɛŋ^{3}	(tœj^{6})	lɔk^{7}
FK	(eŋ^{1})	tsɐm^{1} pan^{3}	–	k^{wh}aj$^{\text{-}33}$ tsuə^{3}	!nœk^{9}
GZ	tsœn$^{1\text{-}55}$	tsɐm^{1} pan^{3}	–	faj^{5} tsi^{3}	kɔk^{9} lɔk^{7} t^{h}ɐw^{3}
NH	tsœn^{1}	tsɐm^{1} paŋ^{3}	–	wa^{5} ty^{3}	kɔk^{9} lɔk^{7} t^{h}aw^{2}
DG	tsiŋ$^{\text{-}55}$	tsɐm^{1} pɛŋ^{3}	–	faj^{5} tsɐj^{3}	kɔk^{8} ŋɔk^{7} t^{h}aw^{3}
FG	tsɐn^{1}	tsɐm^{1} pan^{3}	–	faj^{1} tsy^{3}	kɔk^{9} lɔk^{8} t^{h}ɐw^{2}
YF	tsɐn^{1}	tsɐm^{1} paŋ^{3}	–	faj^{5} tsi^{3}	kɔk^{9}
XH	tsæn^{1}	tsæm^{1} pan^{3}	–	faj^{1} tsi^{3}	kɔk^{9} hæw^{2}
TS	tun$^{\text{-}35}$	–	im^{6} pan^{3}	faj^{1} tu^{3}	kɔk^{8} lɔk^{8} sɔk^{8}
KP	tun$^{\text{-}35}$	–	em^{6}	faj$^{\text{-}35}$	kɔk$^{\text{-}11}$ lɔk$^{\text{-}11}$
EP	tsun$^{\text{-}35}$	–	tiam^{6}	fuaj^{1} tsu^{3}	kuak^{8} luak^{8} suak^{7}
BL	(p^{h}iŋ^{2})	tʃɐm^{1} pan^{3}	–	waj^{5} ti^{3}	kɔk^{9} lɔk^{9}
QZ	(fu^{2})	tʃɐm^{1} pan^{3}	–	faj^{5} tʃi^{3}	kɔk^{9} lɔk^{8} t^{h}ɐw^{2}
NN_p	tʃɐn^{1}	tʃɐm^{1} pan^{3}	–	k^{wh}aj^{5} tʃuj^{6}	kak^{9} tɐw^{2}
BS_p	(ɛŋ^{1})	tʃɐm^{1} pan^{3}	–	k^{wh}aj^{5} tʃuj^{6}	kak^{9} tɐw^{2}
BY	(piŋ^{2})	tʃɐm^{1} pan^{3}	–	(tʃuj^{5})	tʃak^{9} lœk^{10}
MS	(piŋ^{2})	tʃɐm^{1} pan^{3}	–	k^{h}uaj^{5} ɬi^{3}	!nua^{5}
YJ	tʃun^{1}	–	–	k^{h}aj^{5}	–
WC	tun^{1}	–	–	faj^{5} tʃ^{h}i^{6}	kak^{9} t^{h}ɐw^{2}

Table C.17

Gloss	door	passive ptcl.	pincers	salt	soup spoon
Index	86	87	88	89	90
Orth	門	畀	鉗	鹽	匙羹
CY	*muɔn^{2}	*pi^{3}	*gɛm^{2}	jam^{2}	*[ʒ]i^{2} kaŋ^{1}
YS	mon^{-33} kew$^{1-52}$	pɐj^{3}	kin^{2}	im^{2}	tsi^{-33} keŋ$^{1-52}$
LX	mun^{6} kin^{6}	(pɐn^{1})	kin^{2}	jin^{2}	ʃi^{6} kiaŋ^{1}
HJ	mun^{2} kim^{6}	pɐj^{3}	kin^{2}	jin^{2}	tsi^{2} kɛŋ^{1}
FK	mɐn^{-22} kaw^{1}	pej^{3}	kim^{2}	im^{2}	tsi^{-22} keŋ^{1}
GZ	mun^{2} ts^{h}am^{4}	pij^{3}	k^{h}im$^{2-35}$	jim^{2}	ts^{h}i^{2} kɐŋ$^{1-55}$
NH	muk^{8} tɐj^{6} k^{h}am^{3}	pɐj^{3}	k^{h}im^{2}	him^{2}	ts^{h}i^{2} kɐn^{1}
DG	(tɐj^{5} p^{h}uk^{8})	pɐj^{3}	k^{h}im$^{2-35}$	jin^{2}	ts^{h}i^{2} kɛŋ$^{-55}$
FG	(tɐj^{6} puk^{7})	pɐj^{3}	k^{h}im^{2} (tsɐj^{3})	im^{2}	ts^{h}i^{6} kɐŋ^{1}
YF	mun^{2} kaw^{1}	pij^{3}	k^{h}iŋ^{2} (tsɐj^{3})	jin^{2}	ts^{h}i^{2} kaŋ^{1}
XH	mun^{2} k^{h}æm^{4}	pij^{3}	k^{h}im^{2}	(siɔŋ^{6} mij^{6})	ts^{h}i^{2} kaŋ^{1}
TS	(i^{6} fuk^{6})	(i^{3})	k^{h}iam$^{2-35}$	(siaŋ^{6} mi^{-35})	si^{2} kaŋ$^{-35}$
KP	mun^{2} ij^{6} fuk^{7}	(ij^{3})	k^{h}im$^{2-35}$	(siaŋ^{6} mij^{6})	si^{2} kaŋ$^{-35}$
EP	m(ɔ)n^{2} taŋ^{3}	pi^{3}	k^{h}iam$^{2-35}$	(siɔŋ^{6} mi^{-35})	si^{2} kaŋ$^{-35}$
BL	mun^{2} hɐm^{3}	(tʃip^{7})	k^{h}ɛm^{2}	jim^{2}	ʃi^{2} kɛŋ^{1}
QZ	mun^{2} k^{h}ɐm^{3}	(aj^{6})	kɛm^{2}	jim^{2}	(t^{h}ɔŋ^{1} kaŋ^{1})
NN_p	mun^{2} k^{h}am^{3}	(ŋaj^{2})	kɛm^{2}	him^{2}	(tuŋ^{2} kɛŋ^{1})
BS_p	mun^{2} k^{h}am^{3}	(ŋaj^{1})	kɛm^{2}	him^{2}	(tiw^{2} kɛŋ^{1})
BY	mun^{2} k^{h}am^{3}	(tʃɛk^{10})	kɛm^{2}	jim^{2}	(t^{h}œŋ^{1} kaŋ^{1})
MS	mun^{2} kim^{4}	(pi^{6})	kɛm^{2}	him^{2}	(tiw^{2} kɛŋ^{1})
YJ	mun^{2}	pij^{3}	k^{h}iɛm^{2} tʃɐj^{3}	jim^{2}	–
WC	mun^{2}	bi^{3}	k^{h}ɛm^{2}	–	–

Table C.18

Gloss	steamed rice	thief	tile	umbrella	wood stake
Index	91	92	93	94	95
Orth	飯	賊	瓦	遮	木樁
CY	*van^{6}	*[dz]ɐk^{8}	*ʔŋʷa^{3}	*tʃia^{1}	*muk^{8} tʃucŋ^{1}
YS	fan^{-22}	(pa^{-33} sɐw^{3})	ŋa^{-24}	tse$^{1-52}$	(mok^{8} toj^{5})
LX	!pan^{6}	(pa^{2} ʃɐw^{3})	ŋa^{4} pʰɛn^{5}	tʃɛ^{1}	mɔk^{8} tʃɔŋ^{1}
HJ	fan^{6}	tsʰak^{8} tsaj^{3}	ŋa^{4}	tsɛ^{1}	–
FK	!pan^{6}	tsɐk^{8} tsɐj^{3}	ŋa^{4}	tsiə^{1}	(tœ^{6})
GZ	fan^{6}	tsʰak^{8} tsɐj^{3}	ŋa^{4}	tsɛ$^{1-55}$	muk^{8} tsɔŋ$^{1-53}$
NH	pʰaŋ^{6}	(si^{3} mɔ^{1} ta^{3})	ŋa^{4}	tsɛ^{1}	tsɔŋ^{1}
DG	fɛŋ^{5}	tsʰak^{8} tsɔj^{3}	ŋa^{4}	tsø$^{-55}$	muk^{8} tsuŋ^{1}
FG	fan^{6}	tsʰɐk^{8} tsɐj^{3}	ŋa^{3}	tse^{1}	(tak^{9})
YF	faŋ^{2}	(syt^{9} tsɐj^{3})	ŋa^{4}	tsɛ^{1}	tsɔŋ^{1}
XH	fan^{6}	(siw^{3} mɔ^{3})	ŋa^{3} pʰin^{3}	tsæ^{1}	muk^{8} tsɔŋ^{1}
TS	fan^{6}	tʰak^{8} tɔj^{3}	ŋa^{3} pʰen^{3}	tsia^{-35}	muk^{4} tsɔŋ^{1}
KP	fan^{6}	(ɬiw^{3} mu^{3})	ŋa^{3} hin^{3}	tsia^{-35}	muk^{8} tsɔŋ^{1}
EP	fan^{6}	(siw^{3} m(ɔ)^{3} tsaj^{3})	ŋa^{3} pʰi(e)n^{3}	tsia^{-35}	(muk^{8} hak^{8})
BL	fan^{6}	tʰɐk^{8}	ŋa^{4}	tʃɛ^{1}	muk^{8} tʃɔŋ^{1}
QZ	fan^{6}	(pak^{8} nɛm^{1})	ŋa^{4}	tʃɛ^{1}	muk^{8} tʃɔŋ^{1}
NN_p	fan^{6}	tʃʰɐk^{8} tʃaj^{3}	ŋʷa^{4}	(ɬan^{5})	muk^{10} tʃuŋ^{1}
BS_p	fan^{6}	(pa^{2} ʃɐw^{3})	ŋʷa^{4}	(ɬan^{5})	mɔk^{4} tʃuŋ^{1}
BY	fan^{6}	tʃɐk^{8}	ŋʷa^{4}	(ɬan^{5})	muk^{8} tʃœŋ^{1}
MS	!pan^{6}	ɬɐk^{8}	ŋua^{4} pʰin^{5}	tʃɛ^{1}	muk^{8} tʃuaŋ^{1}
YJ	fan^{6}	tʃʰɐk^{8} tʃɐj^{3}	ŋa^{3}	tʃiɛ^{1}	–
WC	fan^{6}	tʰɐt^{8}	ŋa^{4}	–	–

Table C.19

Gloss	blow (wind)	cloud	dust	fire	moon
Index	96	97	98	99	100
Orth	風	雲	塵	火	月
CY	*[xʷ]a[n]¹	*wun²	*dʒin²	*xʷɔ³	*ŋiɔt⁸
YS	(kʷet⁻³³ foŋ¹⁻⁵²)	wɐn²	!fuj¹⁻⁵² tɐn²	tʰaw⁻³³ fɔ³	yt⁸ kʷɔŋ¹⁻⁵²
LX	(kʷɛt⁹ fɔŋ¹)	wɐn²	!fuj⁻²² tɐn²	tin⁻²² fɔ³	ŋyt⁸ kʷɔŋ¹
HJ	(tsʰɔj¹ fuŋ¹)	wɐn²	!fuj¹ tɐn²	hi³ fɔ³	jyt⁸ lœŋ⁶
FK	fat⁻³³ fɔŋ¹	wɐn²	!nɐj⁻²² tɐn²	tem⁻³³ fuə³	ɲyt⁸ lyŋ⁶
GZ	fan¹⁻⁵⁵ fuŋ¹⁻⁵³	wɐn²	jin⁵ tsʰɐn²	fɔ³ mij⁴	jyt⁸ kʷɔŋ¹⁻⁵⁵
NH	waŋ¹ pʰuŋ¹	wɐn²	tsʰɐn² hɔ¹	fɔ³ mɐj⁴	jyt⁸ kʷɔŋ¹
DG	fɛŋ¹ fuŋ¹	vɐn² tsʰuj³	fuj¹ tsʰɐn²	fɔ³ zin⁵	zøt⁸ kʷɔŋ¹
FG	(hi³ fuŋ¹)	wɐn²	fuj¹ tsʰɐn²	tɐw¹ fɔ³	yt⁸ na³
YF	(tsʰu¹ taj² fuŋ¹)	wɐn²	pʰœŋ¹ tsʰɐn²	tʰɐw⁵ fɔ³	jyt⁸ kʷɔŋ¹
XH	fan¹ fuŋ¹	væn²	fuj¹ tsʰæn²	fɔ³	ŋit⁸
TS	(a³ fuŋ¹)	vun²	fɔj¹ tsʰin²	fɔ³	ŋut⁸ kɔŋ¹
KP	fan¹ fuŋ¹	vun²	fɔj¹ tsʰiŋ²	fɔ³	(hu² su⁻³⁵)
EP	(ta³ fuŋ¹)	vun²	f(ɔ)i¹ tsʰian²	fua³	(pʰua² sua⁻³⁵)
BL	fat⁹ fuŋ¹	wɐn²	wuj¹ tʃʰɐn²	hu³ mi²	ɲot⁸ lɛŋ⁶
QZ	fan¹ fuŋ¹	wɐn²	fuj¹ tʃʰɐn²	fɔ³ jim²	ɲit⁸ kɔŋ¹
NN_p	(huj³ fuŋ¹)	wun²	pʰɐn⁵ lɐm²	hu³ miw²	wit¹⁰ kuŋ¹
BS_p	(kʷat⁹ fɔŋ¹)	wən² ɲa²	pʰɐn⁻³⁵ pʰa³	hu³ miw²	wit⁸ lɛŋ⁶
BY	(kʷat⁹ fuŋ¹)	wɐn² tʃʰaj³	mɐn¹ tʃən²	həw³ jim⁶	ɲut¹⁰ lɛŋ⁶
MS	fan¹ fuŋ¹	wɐn²	pʰuŋ¹ tʃən²	hu³ miw²	ɲut⁸ liəŋ⁶
YJ	(pʰɔw⁵)	wɐn²	fuj¹ tʃʰɐn²	–	jit⁸ kɔŋ¹
WC	–	wɐn²	jin¹ tʃʰɐn²	fɔ³	wit⁸ lɛŋ⁶

Table C.20

Gloss	mud	rain	silver	snow	star
Index	101	102	103	104	105
Orth	泥	雨	銀	雪	星
CY	*nɐj^{2}	*jy[^{4b}]	*ŋɐn^{2}	*siɔt^{7}	*siaŋ^{1}
YS	naj^{2}	y^{4}	ŋɐn^{2}	syt^{7}	siŋ$^{1\text{-}42}$ tsaj^{3}
LX	naj^{2} tʰɐw^{3}	jy^{4}	taj$^{\text{-}22}$ ŋɐn^{2}	syt^{7}	siŋ^{1}
HJ	naj^{2}	y^{4}	taj^{6} ŋɐn^{2}	θyt^{7}	tʰin^{1} θɛŋ^{1}
FK	nɐj^{2}	y^{4}	taj^{6} ŋɐn^{2}	syt^{7}	tʰen$^{\text{-}33}$ siŋ^{1}
GZ	nɐj^{2}	jy^{4}	ŋɐn^{3}	syt^{9}	sɛŋ$^{1\text{-}55}$
NH	nɐj^{2}	hy^{4}	ŋɐn^{2}	syt^{9}	tʰin^{1} sɛŋ^{1}
DG	nɔj^{2}	ji^{4}	ŋɐn^{2}	søk^{7}	sɛŋ$^{\text{-}55}$
FG	nɐj^{2}	(søj^{3})	ŋɐn^{2}	syt^{9}	θɐŋ^{3}
YF	nɐj^{2}	jy^{4}	ŋɐn^{2} tsɐj^{3}	syt^{9}	siŋ^{1} tsɐj^{3}
XH	næj^{2}	(suj^{3})	ŋæn^{2}	sit^{9}	sæŋ^{1}
TS	naj^{2}	(suj^{3})	ŋan^{2}	ɬut^{7}	ɬen^{1}
KP	naj^{2}	(suj^{3})	ŋan$^{\text{-}35}$	ɬit^{7}	ɬiaŋ$^{\text{-}35}$
EP	naj^{2}	(suj^{3})	ŋan$^{\text{-}35}$	si(e)t^{7}	siaŋ$^{\text{-}35}$
BL	nɐj^{2}	(ʃuj^{3})	ŋɐn^{2}	ɬyt^{7}	tʰin^{1} ɬiŋ^{1}
QZ	nɐj^{2}	ji^{4}	ŋɐn^{2}	ɬit^{9}	ɬiŋ^{1}
NN_p	nɐj^{2}	huj^{4}	ŋɐn^{2}	ɬit^{9}	ɬiŋ^{1} tʃi^{3}
BS_p	nɐj^{2}	huj^{4}	ŋɐn^{2}	ɬit^{9}	ɬən^{1} tʃi^{3}
BY	nɐj^{2}	huj^{4}	ɲɐn^{2}	ɬut^{9}	ɬiŋ^{1} tʃəj^{3}
MS	nɐj^{2}	(ʃuj^{3})	ŋɐn^{2}	ɬut^{7}	ɬiŋ^{1}
YJ	nɐj^{2}	–	ŋɐn^{2}	ɬit^{9}	tʰin^{1} ɬiŋ^{1}
WC	nɐi^{2}	–	–	ɬut^{9}	tʰin^{1} ɬiŋ^{1}

Table C.21

Gloss	stone	sun	thunder	to fall (of rain)	water
Index	106	107	108	109	110
Orth	石頭	日頭	雷	落	水
CY	*ʒiak^{8} dɐw^{2}	*ɲi(a)t^{8} dɐw^{2}	*luɔj^{2}	*lɔk^{8}	*ʃuj^{3}
YS	sik^{8} taw^{2}	it^{8} taw^{2}	hœŋ^{3} loj^{-33} koŋ$^{1-52}$	lɔk^{-22}	sy^{3}
LX	ʃik^{8} taw^{2}	ɲit^{8} taw^{2}	hyŋ^{3} lœj^{2}	lœk^{8}	ʃy^{3}
HJ	tsɛk^{8} taw^{2}	ɲit^{8} taw^{2}	hœŋ^{3} lɔj^{2} kuŋ^{1}	lɔk^{9}	θuj^{3}
FK	tsik^{-22} tɐw^{2}	nit^{8} tɐw^{2}	hyŋ^{3} lœ^{2} kɔŋ^{1}	lœk^{8}	sœ^{3}
GZ	sɛk^{8}	jit^{8} tʰɐw^{-35}	haŋ^{2} lœj^{2}	lɔk^{8}	sœj^{3}
NH	sɛk^{8} taw^{2}	jit^{8} tʰaw^{2}	haŋ^{2} ly^{2}	lɔk^{8}	sɔ^{3}
DG	søk^{8} tʰaw^{2}	zit^{8} tʰaw^{2}	hɛŋ^{2} ŋuj^{2}	ŋɔk^{8}	suj^{3}
FG	sik^{8} tʰɐw^{2}	jat^{8} tʰɐw^{-35}	ta^{3} løj^{2}	lɔk^{8}	søj^{3}
YF	sɛk^{8} tʰɐw^{2}	jit^{8} tʰɐw^{2}	hœŋ^{3} lu^{2}	lɔk^{8}	su^{3}
XH	sæk^{8} hæw^{2}	ŋit^{8} hæw^{2}	haŋ^{2} luj^{2}	lɔk^{8}	suj^{3}
TS	siak^{-11} ku^{3}	ŋit^{8} hew^{2}	haŋ^{2} luj^{2}	lɔk^{8}	suj^{3}
KP	siak^{8} haw^{2}	ŋit^{8} haw^{2}	luj^{4} hiaŋ^{3}	lɔk^{8}	suj^{3}
EP	siak^{8} hij^{2}	ŋit^{8} hij^{2}	ta^{3} luj^{2}	luak^{8}	suj^{3}
BL	ʃik^{8} kʰɐw^{2}	ɲɐt^{8} tʰɐw^{2}	luj^{2} kuŋ^{1} hɛŋ^{3}	lɔk^{8}	ʃuj^{3}
QZ	ʃɛk^{8} kʰɐw^{-32}	ɲɐt^{8} tʰɐw^{2}	ta^{3} luj^{2}	lɔk^{8}	ɬuj^{3}
NN_p	ɬik^{8} tɐw^{2}	ɲɐt^{10} tɐw^{2}	lɔj^{2} kuŋ^{1} ɬaŋ^{3}	lak^{10}	ɬuj^{3}
BS_p	ʃət^{8} tɐw^{2}	ɲit^{8} tɐw^{2}	ta^{3} lɔj^{2}	luk^{8}	ʃuj^{3}
BY	ʃik^{8} tɐw^{2}	ɲit^{10} tɐw^{2}	taŋ^{3} luj^{2}	lœk^{10}	ʃœ^{3}
MS	tʃik^{8} tɐw^{2}	ɲət^{8} tɐw^{2}	hiəŋ^{3} luj^{2}	luak^{8}	ʃuj^{3}
YJ	ʃik^{8} tʰɐw^{2}	jɐt^{8} tʰɐw^{2}	luj^{2} kuŋ^{1}	lɔk^{8}	–
WC	–	ɲit^{8} tɐw^{2}	(a^{5} kuŋ^{1})	lɔk^{8}	–

Table C.22

Gloss	wind	1	2	3	4
Index	111	112	113	114	115
Orth	風	一	二	三	四
CY	*fuŋ^{1}	*it^{7}	*ɲi^{6}	*sam^{1}	*si^{5}
YS	k^{w}et^{-33} foŋ$^{1-52}$	jɐt^{7}	ji^{6}	sam^{1}	sɐj^{5}
LX	k^{w}ɛt^{9} fɔŋ^{1}	jɐt^{7}	ɲi^{6}	san^{1}	sɐj^{5}
HJ	ts^{h}ɔj^{1} fuŋ^{1}	jɐt^{7}	ɲi^{6}	θam^{1}	θɐj^{5}
FK	fat^{-33} fɔŋ^{1}	jɐt^{7}	ɲ[^{j}]i^{6}	sam^{1}	sɐj^{5}
GZ	fan$^{1-55}$ fuŋ$^{1-53}$	jɐt^{7}	ji^{6}	sam^{1}	sij^{5}
NH	waŋ^{1} p^{h}uŋ^{1}	jɐt^{7}	ji^{6}	sam^{1}	sɐj^{5}
DG	fɛŋ^{1} fuŋ^{1}	zɐk^{7}	zi^{5}	saŋ^{1}	sɐj^{5}
FG	hi^{3} fuŋ^{1}	jɐt^{7}	ji^{6}	θam^{1}	θi^{1}
YF	ts^{h}u^{1} taj^{2} fuŋ^{1}	jɐt^{7}	ji^{2}	sam^{1}	si^{5}
XH	fan^{1} fuŋ^{1}	zæt^{7}	ŋi^{6}	sam	sij^{1}
TS	a^{3} fuŋ^{1}	zit^{7}	ŋij^{6}	ɬam^{1}	ɬ[e]i^{1}
KP	fan^{1} fuŋ^{1}	zit^{7}	ŋij^{6}	ɬam^{1}	ɬij^{1}
EP	ta^{3} fuŋ^{1}	zi[e]t^{7}	ŋi^{6}	sam^{1}	si^{1}
BL	fat^{9} fuŋ^{1}	jɐt^{7}	ɲi^{6}	ɬam^{1}	ɬi^{5}
QZ	fan^{1} fuŋ^{1}	jɐt^{7}	ɲi^{6}	ɬam^{1}	ɬi^{5}
NN_p	huj^{3} fuŋ^{1}	ɐt^{9}	ɲi^{6}	ɬam^{1}	ɬuj^{5}
BS_p	k^{w}at^{9} fɔŋ^{1}	ɐt^{9}	ɲi^{6}	ɬam^{1}	ɬuj^{5}
BY	k^{w}at^{9} fuŋ^{1}	jət^{7}	ɲi^{6}	ɬam^{1}	ɬəj^{5}
MS	fan^{1} fuŋ^{1}	jət^{7}	ŋi^{6}	ɬam^{1}	ɬi^{5}
YJ	fuŋ^{1}	jɐt^{7}	ji^{6}	ɬam^{1}	ɬij^{5}
WC	fuŋ^{1}	jɐt^{7}	ɲi^{6}	ɬam^{1}	ɬi^{5}

Table C.23

Gloss	5	6	7	8	9
Index	116	117	118	119	120
Orth	五	六	七	八	九
CY	*ʔŋu^{3}	*luk^{8}	*tsʰit^{7}	*pat^{7}	*kiw^{3}
YS	m̩^{4}	lok^{8}	tsʰɐt^{7}	pet^{9}	kɐw^{3}
LX	m̩^{4}	lɔk^{8}	tsʰɐt^{7}	pɛt^{9}	kɐw^{3}
HJ	ŋ̍^{4}	luk^{8}	tsʰɐt^{7}	pɛt^{9}	tsaw^{3}
FK	ŋ̍^{4}	lɔk^{8}	tsʰɐt^{7}	pat^{9}	tsɔw^{3}
GZ	ŋ̍^{4}	luk^{8}	tsʰɐt^{7}	pat^{9}	kɐw^{3}
NH	ŋ̍^{4}	luk^{8}	tʰɐt^{7}	pik^{9}	kɐw^{3}
DG	m̩^{4}	ŋuk^{8}	tsʰɐk^{7}	pɛ^{9}	kaw^{3}
FG	ŋ̍^{3}	luk^{8}	tsʰɐt^{7}	pat^{9}	kɐw^{3}
YF	ŋ̍^{4}	luk^{8}	tsʰɐt^{7}	pak^{9}	kɐw^{3}
XH	ŋ̍^{3}	luk^{8}	tsʰæt^{7}	pat^{9}	kæw^{3}
TS	m̩^{3}	luk^{8}	tʰit^{7}	pat^{9}	kiw^{3}
KP	m̩^{3}	luk^{8}	tʰik^{7}	vat^{9}	kiaw^{3}
EP	m̩^{3}	luk^{8}	tsʰiat^{7}	pat^{9}	kij^{3}
BL	ŋ̍^{4}	luk^{8}	tʰɐt^{7}	pat^{9}	kɐw^{3}
QZ	ŋ̍^{4}	luk^{8}	tʃʰɐt^{7}	pat^{9}	kɐw^{3}
NN_p	ŋɔ^{4}	luk^{10}	tʃʰɐt^{9}	pat^{9}	kɐw^{3}
BS_p	ŋɔ^{4}	lɔk^{4}	tʃʰɐt^{9}	pat^{9}	kɐw^{3}
BY	ŋəw^{4}	luk^{8}	tʃʰɐt^{7}	pat^{9}	tʃəw^{3}
MS	ŋ̍^{4}	luk^{8}	tʰɐt^{7}	pat^{7}	tʃɐw^{3}
YJ	wuŋ^{3}	luk^{8}	tʃʰɐt^{7}	pat^{9}	kiɛw^{3}
WC	ŋu^{4}	luk^{8}	tʰɐt^{7}	bat^{9}	kɐw^{3}

Table C.24

Gloss	10	100	1000	10000	general mw
Index	121	122	123	124	125
Orth	十	百	千	萬	個
CY	*ʒ’ ip^{8}	*pak^{7}	*ts^{h}ian^{1}	*man^{6}	*kɔ^{5}
YS	sɐp^{8}	pek^{9}	ts^{h}in^{1}	man^{6}	(la$^{1\text{-}52}$)
LX	ʃɐt^{8}	piak^{9}	ts^{h}in^{1}	man^{6}	(na^{1})
HJ	θɐp^{8}	pɛk^{9}	ts^{h}in^{1}	man^{6}	kɔ^{5}
FK	sɐp^{8}	pek^{9}	ts^{h}en^{1}	man^{6}	kuə^{5}
GZ	sɐp^{8}	pak^{9}	ts^{h}in^{1}	man^{6}	kɔ^{5}
NH	sɐp^{8}	pak^{9}	t^{h}in^{1}	maŋ^{6}	kɔ^{5}
DG	sɔk^{8}	pɛ^{9}	ts^{h}in^{1}	mɛŋ^{4}	kɔ^{5}
FG	sɐp^{8}	pak^{9}	ts^{h}in^{1}	man^{6}	kɔ^{1}
YF	sɐp^{8}	pak^{9}	ts^{h}in^{1}	maŋ^{2}	kɔ^{5}
XH	sæp^{8}	pak^{9}	ts^{h}in^{1}	man^{6}	kɔ^{1}
TS	sip^{8}	pak^{7}	t^{h}en^{1}	man^{6}	(kɔj^{1})
KP	sip^{8}	vak^{7}	t^{h}in^{1}	man^{6}	(kɔj^{1})
EP	si[ə]p^{8}	pak^{9}	ts^{h}ian^{1}	man^{6}	kua^{1}
BL	ʃɐp^{8}	pik^{9}	t^{h}in^{1}	man^{6}	(tʃik^{7})
QZ	ʃɐp^{8}	pak^{9}	tʃ^{h}in^{1}	man^{6}	(tʃɛk^{9})
NN_p	ɬɐp^{8}	pɛk^{9}	tʃ^{h}in^{1}	man^{6}	(ki^{5})
BS_p	ʃɐp^{8}	pɛk^{9}	tʃ^{h}in^{1}	man^{6}	(ki^{5})
BY	ʃɐp^{8}	pak^{9}	tʃ^{h}in^{1}	man^{6}	kœ^{5}
MS	ʃɐp^{8}	pɛk^{55}	t^{h}in^{1}	wuan^{6}	kɔ^{5}
YJ	ʃɐp^{5}	pak^{9}	tʃ^{h}in^{1}	man^{6}	kɔ^{6}
WC	ʃɐp^{8}	bak^{9}	t^{h}in^{1}	man^{6}	(tʃik^{7})

Table C.25

Gloss	instance cntr	mw for house	China fir	amaranth (1)	amaranth (2)
Index	126	127	128	129	130
Orth	次	間	杉樹	莧菜	莧菜
CY	*tsʰʐ̩^{5}	*kan^{1}	*tʃʰam^{5} ʒy^{6}	*ɣian^{6} tsʰɔj^{5}	*ɣan^{6} tsʰɔj^{5}
YS	tsʰɐj^{5}	(tsɔ^{6})	tsʰam^{5} tsy^{6}	hin^{6} tsʰɔj^{5}	–
LX	tsʰɐj^{5}	kɛn^{-33}	(ʃa^{-22} tʃy^{6})	–	hɛn^{-22} tsʰœj^{5}
HJ	tsʰɐj^{5}	kɛn^{1}	tsʰam^{5}	–	hen^{6} tsʰoj^{5}
FK	tsʰi^{5}	kan^{1}	tsʰam^{5}	jin^{6} tsʰɔj^{5}	–
GZ	tsʰi^{5}	kan^{1}	tsʰam^{5} sy^{6}	jin^{6} tsʰɔj^{5}	–
NH	(wɐn^{2})	kɛŋ^{1}	tsʰam^{5} tsʰy^{6}	–	!hɛm^{6} tsʰɔ^{5}
DG	tsʰɐj^{5}	kɛŋ^{1}	tsʰaŋ$^{1-35}$ sy^{5}	hin^{5} tsʰuj^{5}	–
FG	tsʰy^{1}	kan^{1}	tsʰam^{1} sy^{6}	–	!hɔn^{6} tsʰøj^{5}
YF	tsʰi^{5}	kaŋ^{1}	tsʰam^{1} sy^{2}	hin^{6} tsʰu$^{5-35}$	–
XH	tsʰi^{1}	kæn^{1}	tsʰam^{1} si^{6}	–	han^{6} tsʰuj^{1}
TS	(lun^{2})	kan^{1}	tsʰem^{1} su^{6}	–	han^{6} tʰɔj$^{1-21}$
KP	(lun^{6})	kan^{1}	tsʰam$^{1-35}$ si^{6}	–	han^{6} tʰɔj$^{1-31}$
EP	su^{1}	kan^{1}	tsʰiam^{-35} si^{6}	–	han^{6} tsʰɔj$^{1-22}$
BL	(kyn^{2})	kan^{1}	tʃʰam^{5} muk^{8}	–	han^{6} tʰoj^{5}
QZ	tʃʰi^{5}	kan^{1}	tʃʰam^{5} muk^{8}	–	han^{6} tʃʰɔj^{5}
NN_p	(kʷɐn^{5})	kan^{1}	(ɬa^{1} muk^{10})	–	han^{6} tʃʰaj^{5}
BS_p	(tʰaŋ^{5})	(ɔk^{9})	(ɬa^{1} ʃuj^{6})	–	han^{6} tʃʰaj^{5}
BY	tʃʰəj^{5}	(tʃəw^{6})	tʃʰam^{1} muk^{8}	–	han^{6} tʃʰaj^{5}
MS	(tʰuaŋ^{5})	(ɬɔ^{6})	(ʃa^{1} muk^{8})	–	han^{6} tʰɔj^{5}
YJ	tʃʰij^{5}	kan^{1}	tʃʰam^{5} ʃi^{6}	–	han^{6} tʃʰɔj^{5}
WC	–	(ŋan^{3})	–	jin^{6} tʰɔj^{5}	–

Table C.26

Gloss	apricot	bamboo	bamboo sprout	banana	barnyard grass
Index	131	132	133	134	135
Orth	杏	竹	茭筍	蕉	稗
CY	ɣɐŋ^{6}	*tʃuk^{7}	*kaw^{1} syn^{3}	*tsiaw^{1}	*bɐj^{6}
YS	heŋ^{6}	!tok^{7}	wɔ^{2} sen^{3}	hœŋ^{1} tsiw^{1}	paj^{6}
LX	haŋ^{2}	!tɔk^{-22} tsɐj^{3}	kɛw^{-22} ʃɐn^{3}	hyŋ$^{-22}$ tsiw^{1}	paj^{-22} tsɐj^{3}
HJ	hɛŋ^{6}	!tuk^{7}	tuŋ^{1} sɐn^{3}	tsiw^{1} tsaj^{3}	(jɛ^{4} wɔ^{2})
FK	eŋ^{6}	tsɔk^{7}	–	hyŋ^{1} tsew^{1}	–
GZ	hɐŋ^{6}	tsuk^{7}	kaw^{1} sœn$^{1-35}$	tsiw^{1}	pɐj^{-35}
NH	hɐn^{6}	tsuk^{7}	kɛw^{1} sœn$^{1-35}$	tiw^{1}	pɐj^{6}
DG	hɐn^{5}	tsuk^{7}	kaw^{1} søn$^{3-35}$	tsiw$^{1-55}$	pɔj$^{6-35}$
FG	hɐŋ^{6}	tsuk^{7}	kaw^{1} ts^{h}ɐn$^{2-35}$	tsiw^{1}	pɐj^{6} tsɐj^{3}
YF	hɐn^{2}	tsuk^{7}	–	hœŋ^{1} tsiw^{1}	pɐj^{2} tsɐj^{3}
XH	haŋ^{6}	tsuk^{7}	kaw^{1} sæn^{1}	tsiw^{1}	p^{h}aj^{6}
TS	haŋ^{6}	tsuk^{9}	kaw^{1} ɬun^{2}	hiaŋ^{1} tiaw^{-21}	p^{h}aj^{-21}
KP	haŋ^{4}	tsuk^{7}	kaw^{1} ɬun^{2}	hiaŋ^{1} tiw^{-21}	haj^{-21}
EP	haŋ$^{6-35}$	tsuk^{7}	kaw^{1} sun^{1}	tsiw^{-31} kua^{3}	p^{h}aj^{3}
BL	hɐŋ^{6}	tʃuk^{7}	kaw^{1} pɛk^{-22}	tʃiw^{1} ti^{3}	paj^{6} ti^{3}
QZ	hɐŋ^{6}	tʃuk^{7}	kaw^{1} ɬɐn^{3}	pa^{1} tʃiw^{1}	paj^{6} tʃ^{h}ɐw^{3}
NN_p	hɛŋ^{6}	tʃuk^{9}	kaw^{1} ɬɐn^{3}	tʃiw^{1} tʃi^{3}	(wɐŋ^{1})
BS_p	hən^{5}	tʃɔk^{9}	kaw^{1} ɬən^{3}	tʃiw^{1} tʃi^{3}	(wɐŋ^{1})
BY	haŋ^{6}	tʃuk^{7}	kaw^{1} pak^{8}	pa^{1} tʃiw^{1}	paj^{6} tʃəj^{3}
MS	hiŋ^{5}	tʃuk^{7}	tuŋ^{1} ɬɐn^{3}	ɬiw^{1} ɬi^{3}	paj^{6} ɬi^{3}
YJ	haŋ^{6} tʃij^{3}	tʃuk^{7}	–	tʃiw^{1} tʃij^{3}	pɐj^{6}
WC	–	tʃuk^{7}	–	tiw^{1} ti^{3}	–

Table C.27

Gloss	black jujube	chestnut	chinaberry tree	coriander (1)	coriander (2)
Index	136	137	138	139	140
Orth	黑棗	栗	苦楝	芫荽	芫荽
CY	*x[ɐ]k^{7} tsɔw^{3}	*lyt^{8}	*xu^{3} lian^{6}	*jɔm^{2} s[ɐ]j^{1}	*jam^{2} sɐj^{1}
YS	hɐk^{7} tso^{3}	foŋ^{1} lɐt^{8}	fu^{3} len^{6} tsy^{6}	jyn^{2} saj^{1}	–
LX	wu^{-22} tsaw^{3}	pan^{5} lɐt^{8}	fu^{-35} lɛn^{6}	–	jin^{-22} saj^{1}
HJ	hak^{7} tsɔw^{3}	lɐt^{8}	!fɔ^{3} lin^{6} tsɔw^{6}	wɐn^{2} θaj^{1}	–
FK	hɐk^{7} tsɔ^{3}	lɐt^{8}	fu^{3} lem^{6}	jyn^{2} sɐj^{1}	–
GZ	hak^{7} tsuw^{3}	fuŋ^{1} lœt$^{8-35}$	(sɐm^{1} sy^{6})	jyn^{2} sɐj^{1}	jim^{2} sɐj^{1}
NH	hak^{7} tɔ^{3}	p^{h}uŋ^{1} lœt$^{8-35}$	(sɐm^{1} ts^{h}y^{6})	hyn^{2} sɐj^{1}	–
DG	hak^{7} tsɔw^{3}	ŋɐk^{8} mɔj^{4}	!fɔ$^{-55}$ ŋin^{5} sy^{5}	–	zin^{2} sɔj^{-55}
FG	hɐk^{7} tsaw^{3}	fuŋ^{1} lɐt^{8}	fu^{3} lin^{6}	–	jim^{2} θɐj^{1}
YF	hak^{7} tsuw^{3}	lɐt^{8} tsi^{3}	fu^{3} nin^{1} sɐm^{1}	jyn^{2} sɐj^{1}	–
XH	hak^{7} tsu^{3}	læt^{8}	fu^{3} lin$^{6-35}$	zin^{2} sui^{1}	–
TS	hak^{7} taw^{3}	lut^{8} t[ɤ]^{3}	fu^{3} len^{6}	z[ua]n^{2} ɬuj^{-35}	–
KP	hak^{7} tɔ^{3}	lut^{8} e^{5}	fu^{3} liak^{8} e^{5}	zuan^{2} ɬaj^{-35}	–
EP	hak^{7} tsaw^{3}	lut$^{8-35}$	fu^{3} lian$^{6-35}$	–	zi[e]m^{2} saj^{-35}
BL	hɐk^{7} taw^{3}	lɐt^{8} ti^{3}	fu^{3} lin^{6}	–	jim^{2} ɬɐj^{1}
QZ	hak^{7} tʃɐw^{3}	pan^{3} lɐt^{8}	fu^{3} lin^{6}	–	jim^{2} ɬɐj^{1}
NN_p	hɐk^{9} tʃaw^{3}	pan^{3} lɐt^{10}	hɔ^{3} lin^{6}	–	him^{2} ɬɐj^{1}
BS_p	hɐk^{9} tʃaw^{3}	pan^{3} lɐt^{4}	hɔ^{3} lin^{6}	–	him^{2} ɬɐj^{1}
BY	hɐk^{7} tʃœ^{3}	pan^{3} lət^{8}	həw^{3} lin^{6}	–	jim^{2} ɬɐj^{1}
MS	hɐk^{7} ɬəw^{3}	pan^{3} lɐt^{8}	hu^{3} lin^{6}	jun^{2} ɬɐj^{1}	–
YJ	hɐk^{7} tʃɔw^{3}	lɐt^{8} tʃij^{3}	fu^{3} lin^{6} ʃi^{6}	jin^{2} ɬɐj^{1}	–
WC	–	–	!fɔ^{3} lin^{6} muk^{8}	–	–

Table C.28

Gloss	corn	dried persimmon	eggplant (1)	eggplant (2)	fruit
Index	141	142	143	144	145
Orth	粟	柿餅	矮瓜	茄子	生果
CY	*suk^{7}	*ʒ[ʐ̩][^{4b}] piaŋ^{3}	*ɐj^{3} k^{w}a^{1}	*giɔ^{2}	*k^{w}ɔ^{3}
YS	pew^{1} mek^{8}	kiŋ^{1} si^{6}	aj^{3} k^{w}a^{1}	–	seŋ^{1} k^{w}ɔ^{3}
LX	pɛw^{-22} miak^{8}	ʃi^{-22} piŋ^{3}	–	kɛ$^{-22}$ tsɐj^{3}	k^{w}ɔ^{3}
HJ	pɛw^{1} θuk^{7}	tsaj^{6} pɛŋ^{3}	aj^{3} k^{w}a^{1}	–	k^{w}ɔ^{3}
FK	sɔk^{7} mɐj^{4}	tse^{4} piŋ^{3}	aj^{3} k^{w}a^{1}	–	seŋ^{1} kuə^{3}
GZ	suk^{7} mɐj^{4}	ts^{h}i^{-35} pɛŋ^{3}	ɐj^{3} k^{w}a^{1}	–	saŋ^{1} k^{w}ɔ^{3}
NH	suk^{7} mɐj^{4}	ty^{-35} pɛŋ^{3}	ɐj^{3} k^{w}a^{1}	–	saŋ^{1} k^{w}ɔ^{3}
DG	paw^{1} suk^{7}	sɔj^{4} pøŋ^{3}	ŋaj^{3} k^{w}a$^{1-55}$	–	sɛŋ^{1} k^{w}ɔ^{3}
FG	θuk^{7} mɐj^{3}	ts^{h}ɐj^{1} piŋ^{3}	aj^{3} k^{w}a^{1}	k^{h}ø^{2} tsɐj^{3}	saŋ^{1} k^{w}ɔ^{3}
YF	paw^{1} suk^{7}	ts^{h}i^{2} pɛŋ^{3}	aj^{3} k^{w}a^{1}	–	su^{3} ku^{3}
XH	suk^{7} mæj^{3}	si^{6} pæŋ^{3}	aj^{3} k^{w}a^{1}	k^{h}ia^{-35}	saŋ^{1} kɔ^{3}
TS	ɬuk^{7} maj^{3}	si^{4} piaŋ^{3}	–	k^{h}ia^{-35}	saŋ^{1} kɔ^{3}
KP	ɬuk^{7} maj^{3}	ɬu^{4} viaŋ^{3}	k^{h}ia^{-35} ai^{3} ka$^{1-35}$	k^{h}ia^{-35} ai^{3} ka$^{1-35}$	saŋ^{1} kua^{3}
EP	paw^{1} suk^{7}	su^{6} piaŋ^{3}	aj^{3} ka$^{1-35}$	–	saŋ^{1} kua^{3}
BL	paw^{1} ɬuk^{7}	ʃi^{6} piŋ^{3}	aj^{3} k^{w}a^{1}	–	k^{w}ɐt^{8} ti^{3}
QZ	ɬuk^{7} mɐj^{4}	tʃ^{h}i^{5} piɛŋ^{3}	ŋaj^{3} k^{w}a^{1}	–	ʃaŋ^{1} kɔ^{3}
NN_p	(ɲu^{6} mɐj^{4})	ɬɐj^{6} piŋ^{3}	ɐj^{3} k^{w}a^{1}	–	ku^{3} tʃi^{3}
BS_p	(ɲuj^{6} mɐj^{4})	ʃɐj^{6} pən^{3}	(ki^{2} k^{w}a^{1})	–	ʃuj^{3} ku^{3}
BY	(ɬɐt^{9} pœw^{1})	ʃɐj^{4} pəŋ^{3}	(luk^{8} ɬəw^{1})	–	ʃœ^{3} kœ^{3}
MS	(paw^{1} mɐj^{4})	ʃi^{4} piŋ^{3}	aj^{3} kua^{1}	–	kɔ^{3} ɬi^{3}
YJ	paw^{1} ɬuk^{7}	ʃi^{6} piŋ^{3}	–	k^{h}iɛ^{5} tʃij^{3}	ʃuj^{3} kɔ^{3}
WC	tʃɐn^{1} tʃi^{1} ɬuk^{7}	–	–	–	–

Table C.29

Gloss	garlic	garlic chive	ginger	glutinous rice	grain
Index	146	147	148	149	150
Orth	蒜	韭黃	薑	糯米	穀
CY	*suɔn^{5}	*kiw^{3} wɔŋ^{2}	*kiɔŋ^{1}	*nuɔ^{6} ʔmɐj^{3}	*kuk^{7}
YS	syn^{6} taw^{2}	kɐw^{3} ts^{h}oj^{5} wɔŋ^{2}	kœŋ^{1}	nɔ^{6} maj^{4}	kok^{7}
LX	syn^{5}	kɐw^{-33} ts^{h}œj^{5}	kyŋ^{1}	nɔw^{-22} maj^{4}	kɔk^{7}
HJ	θyn^{5} taw^{2}	tsaw^{3} wɔŋ$^{2-35}$	kœŋ^{1}	nɔ^{5} maj^{4}	kuk^{7}
FK	syn^{5}	tsɔw^{3} ts^{h}ɔj^{5}	kyŋ^{1}	nuə^{6} mɐj^{4}	kɔk^{7}
GZ	syn^{5} t^{h}ɐw^{2}	kɐw^{3} wɔŋ^{2}	kœŋ^{1}	nɔ^{6} mɐj^{4}	kuk^{7}
NH	syn^{5} t^{h}aw^{2}	kɐw^{3} fɔŋ^{2}	kœŋ^{1}	nœ^{6} mɐj^{4}	kuk^{9}
DG	søn^{5} tsɐj^{35}	kaw^{3} vɔŋ^{2}	køŋ^{1}	nɔ^{5} mɔj^{4}	kuk^{7}
FG	θyn^{1} tsy^{3}	kɐw^{3} wɔŋ^{2}	køŋ^{1}	nɔ^{6} mɐj$^{3-35}$	kuk^{7}
YF	syn^{5} t^{h}ɐw^{2}	kɐw^{3} wɔŋ^{2}	kœŋ^{1} na^{-35}	nɔ^{2} mɐj^{4}	kuk^{7}
XH	sun^{-35}	kæw^{3} vɔŋ^{2}	kiɔŋ^{1}	nɔ^{6} mæj^{3}	kuk^{7}
TS	ɬɔn^{1}	kiw^{3} vɔŋ^{2}	kiaŋ^{1} na^{3}	nɔ^{6} maj^{3}	kuk^{7}
KP	ɬuan^{1}	kiaw^{3} vɔŋ^{2}	kiaŋ^{1} na^{3}	nu^{6} maj^{3}	kuk^{7}
EP	suan^{-35}	kij^{3} vuaŋ^{2}	kiɔŋ^{1} na^{3}	nua^{6} maj^{3}	kuk^{7}
BL	ɬyn^{5}	tʃɐw^{3} ɔŋ^{2}	kɛŋ^{1}	no^{6} mɐj^{4}	kuk^{7}
QZ	ɬin^{5}	kɐw^{3} wɔŋ^{2}	kiɛŋ^{1}	nɔ^{6} mɐj^{4}	kuk^{7}
NN_p	ɬun^{5}	kɐw^{3} huŋ^{2}	kɛŋ^{1}	nu^{6} mɐj^{4}	kuk^{9}
BS_p	ɬun^{5}	kɐw^{3} huŋ^{2}	kɛŋ^{1}	nu^{6} mɐj^{4}	kɔk^{9}
BY	ɬun^{5}	tʃəw^{3} uŋ^{2}	kɛŋ^{1}	nəw^{6} mɐj^{4}	kuk^{7}
MS	ɬun^{5}	tʃɐw^{3} hɔŋ^{2}	k^{h}iəŋ^{1}	nɔ^{6} mɐj^{4}	kuk^{7}
YJ	ɬun^{5}	!k^{h}iw^{3} tʃ^{h}ɔj^{5}	kiɛŋ^{1}	nɔ^{6} mɐj^{3}	kuk^{7}
WC	ɬun^{5} t^{h}ɐw^{2}	(faj^{5} t^{h}ɔj^{5})	–	!nɛ^{6}	kuk^{7}

Table C.30

Gloss	green onion	growing grain	longan	loquat	lotus flower
Index	151	152	153	154	155
Orth	蔥	禾	龍眼	枇杷	蓮花
CY	*tsʰuŋ¹	*wɔ²	*luŋ² ʔŋan³	*bi² ba²	*lian² xʷa¹
YS	tsʰoŋ¹	wɔ²	loŋ² ŋen⁴	pɐj² pa²	–
LX	tʃʰɔŋ¹	wɔ²	lɔŋ⁻²² ŋɛn⁴	pɐj⁻²² pa²	hɔ⁻²² fa¹
HJ	tsʰuŋ¹	wɔ²	luŋ² ŋɛn⁴	pʰɐj² pʰa² kʷɔ³	–
FK	tsʰɔŋ¹	uə²	lɔŋ² ŋan⁴⁻³⁵	pej² pa²⁻³⁵	–
GZ	tsʰuŋ¹	wɔ²	luŋ² ŋan⁴⁻³⁵	pʰij² pʰa² kʷɔ³	lin² fa¹
NH	suŋ¹	wɔ²	luŋ² ŋɛŋ⁴	(lɔ² kʷɐt⁷)	lin² wa¹
DG	tsʰuŋ¹⁻⁵⁵	vɔ²	ŋuŋ² ŋɛŋ⁴⁻³⁵	pʰɐj² pʰa²	ŋin² fa¹
FG	tsʰuŋ¹	wɔ² tsɐj³	luŋ² ŋan³⁻³⁵	pʰɐj² pʰa²	–
YF	tsʰuŋ¹	(kuk⁷ juŋ¹)	luŋ² ŋaŋ⁴	pʰij² pʰa²	–
XH	tsʰuŋ	vɔ²	luŋ² ŋan³	pʰij² pʰa²	lin² fa¹
TS	tʰuŋ¹⁻³⁵	vɔ²	zɔn² ŋan³	pʰ[e]i² pʰa²	len² ŋɐw⁴ fa¹
KP	tʰuŋ¹⁻³⁵	vua²	(zuŋ² ŋan³)	pʰij² ha²	lin² fa¹
EP	tsʰuŋ¹⁻³⁵	vua²	luŋ² ŋan³	pʰi² pʰa²	lian² ŋij⁶ fa¹
BL	tʰuŋ⁵	(kuk⁷)	luŋ² ŋan⁴	pʰi² pʰa²	lin² wa¹
QZ	tʃʰuŋ¹	wɔ²	luŋ² ŋan⁴	pʰi² pʰa²	lin² ŋɐw⁴ fa¹
NN_p	tʃʰuŋ¹	(kuk⁹)	luŋ² ɲan⁴	puj² pa²	lin² ŋʷa¹
BS_p	tʃʰɔŋ¹	(kɔk⁹)	lɔŋ² ɲan⁴	puj² pa²	ɲɐw⁴ lin² wa¹
BY	tʃʰuŋ¹	(kuk⁷)	luŋ² ɲan⁴	pəj² pa²	hœ² wa¹
MS	tʰuŋ¹	(kuk⁷)	luŋ² ŋan⁴	pi² pa²	lin² ŋɐw⁴ fa¹
YJ	tʃʰuŋ¹	(kuk⁷)	luŋ² ŋan³	pʰij² pʰa²	hɔ² fa¹
WC	ʃaŋ¹ tʰuŋ¹	wɔ²	(lik⁸ ŋan⁴)	–	–

Table C.31

Gloss	lotus root	lotus seeds	lychee	mulberry	nonglutinous rice
Index	156	157	158	159	160
Orth	藕	蓮子	荔枝	桑樹	黏米
CY	*ŋɐw^{4b}	*lian^{2} tsʐ̩^{3}	*lɐj^{6} tʃi^{1}	*ɬɔŋ^{1}	*tʃiam^{1} ʔmɐj^{3}
YS	ŋaw^{4}	lin^{2} tsi^{3}	lɐj^{6} tsi^{1}	sɔŋ^{1}	tsin^{1} maj^{4}
LX	lɛn^{-22} ŋaw^{4}	lɛn^{-22} tsɐj^{3}	lɐj^{-22} tʃi^{1}	sœŋ^{1} tʃy^{6}	(kiaŋ$^{-21}$ maj^{4})
HJ	lin^{2} ŋaw^{4}	lin^{2} tsɐj^{3}	lɐj^{6} tsi^{1}	θɔŋ^{1}	tsin^{1} maj^{4}
FK	len^{2} ŋɐw^{4}	len^{2} tsuə^{3}	lej^{6} tsi^{1}	sœŋ^{1} tsy^{6}	tsim^{1} mɐj^{4}
GZ	lin^{2} ŋɐw^{4}	lin^{2} tsi^{3}	lɐj^{6} tsi^{1}	sɔŋ^{1} sy^{6}	tsim^{1} mɐj^{4}
NH	ŋɐw^{4}	lin^{2} juk^{8}	lɐj^{6} tsi^{1}	sɔŋ^{1} ts^{h}y^{6}	tsim^{1} mɐj^{4}
DG	ŋin^{2} ŋaw^{3}	ŋin^{2} tsɐj^{3}	ŋɔj^{5} tsi^{1}	sɔŋ^{1} sy^{5}	tsin^{1} mɔj^{4}
FG	ŋɐw$^{3\text{-}35}$	lin^{2} tsy^{3}	lɐj^{6} ki^{1}	θɔŋ^{1} sy^{6}	tsim^{1} mɐj$^{3\text{-}35}$
YF	ŋɐw^{4}	lin^{2} tsi^{3}	lɐj^{2} tsi^{1}	sɔŋ^{1} sy^{2}	tsin^{1} mɐj^{4}
XH	ŋæu^{-35}	lin^{2} ti^{3}	laj^{6} tsi^{1}	sɔŋ^{4} si^{6}	tsim^{1} mæj^{3}
TS	ŋew^{-35}	len^{2} tu^{3}	laj^{6} tsi^{1}	ɬɔŋ$^{4\text{-}21}$ su^{6}	tsiam^{1} maj^{3}
KP	lin^{2} ŋaw^{4}	lin^{2} ti^{3}	(kua^{3})	ɬɔŋ$^{1\text{-}21}$ si^{6}	tsi[e]m^{1} maj^{3}
EP	lian^{2} ŋij^{6}	lian^{2} tsu^{3}	!laj^{6} ki^{-35}	suaŋ$^{6\text{-}35}$	tsi[e]m^{1} maj^{3}
BL	ŋɐw^{4}	lin^{2} ti^{3}	lɐj^{6} tʃi^{1}	ɬɔŋ^{1} muk^{8}	tʃim^{1} mɐj^{4}
QZ	ŋɐw^{4}	lin^{2}	li^{6} tʃi^{1}	ɬɔŋ^{1} muk^{8}	(kaŋ^{1} mɐj^{4})
NN_p	ŋɐw^{4}	lin^{2} tʃi^{3}	lɐj^{6} tʃi^{1}	ɬaŋ^{1} muk^{10}	(kɛŋ^{1} mɐj^{4})
BS_p	ɲɐw^{4}	lin^{2} tʃi^{3}	lɐj^{6} tʃi^{1}	ɬaŋ^{1} ʃuj^{6}	(kɛŋ^{1} mɐj^{4})
BY	ŋɐw^{4}	lin^{2} tʃəj^{3}	ləj^{2} tʃi^{1}	ɬœŋ^{1} ʃuj^{6}	(kaŋ^{1} mɐj^{4})
MS	ŋɐw^{4}	lin^{2} ɬi^{3}	li^{6} tʃi^{1}	ɬuaŋ^{1} muk^{8}	(kɛŋ^{1} mɐj^{4})
YJ	ŋɐw^{3}	lin^{2} tʃij^{3}	laj^{6} tʃi^{1}	ʃɔŋ^{1} ʃi^{6}	–
WC	lin^{1} ŋɐw^{4}	–	!lik^{8} tʃi^{1}	–	–

Table C.32

Gloss	olive	peach	pear	persimmon (1)	persimmon (2)
Index	161	162	163	164	165
Orth	欖	桃	梨	柿	柿
CY	*ʔlam^{3}	*dɔw^{2}	*li^{2}	ʒ[ʐ̩]^{6}	ʒ[ʐ̩][^{4b}]
YS	lam^{3}	to^{2} kʷɔ^{3}	sa^{1} lɐj^{2}	tsɐj^{6} kʷɔ^{3}	–
LX	kan^{-22} lan^{3}	tɔw^{-22} tsɐj^{3}	lɐj^{2}	ʃi^{-22} tsɐj^{3}	–
HJ	lam^{3}	tɔw^{2}	lɐj^{2}	tsaj^{3}	–
FK	kuə^{5} lam^{3}	tɔ^{2}	lɐj$^{2-35}$	–	tse^{4}
GZ	pak^{8} lam^{-35}	tʰuw$^{2-35}$	lij$^{2-35}$	tsʰi^{-35}	–
NH	pak^{8} lam^{3}	tʰɔ^{2}	lɐj^{2}	ty^{-35}	–
DG	pɛk^{8} ŋaŋ$^{-35}$	tʰɔw$^{1-35}$	sa^{1} ŋɐj^{2}	–	lɐm^{2} sɔj^{4}
FG	lam^{-35}	tʰaw^{2} tsɐj^{3}	lɐj^{2}	–	tsʰɐj^{-33} tsɐj^{3}
YF	pak^{8} lam^{-35}	tʰuw^{2} tsɐj^{3}	lij^{2}	–	tsʰi^{3} tsɐj^{3}
XH	lam^{3}	hu^{2}	lij^{2}	si^{6}	–
TS	vɔŋ^{2} lam^{3}	haw$^{2-35}$	l[e]i$^{2-35}$	ɬu^{6} tɔt^{7}	–
KP	vɔŋ^{2} lam^{3}	hɔ$^{2-35}$	lij$^{2-35}$	ɬu^{6} tuat^{7}	–
EP	vuaŋ^{2} lam^{3}	haw^{2}	li^{2}	su^{6} tsuat^{7}	–
BL	lam^{3} ti^{3}	tʰɐw^{2} ti^{3}	li^{1} ti^{3}	ʃi^{6} ti^{3}	–
QZ	lam^{3} tʃi^{3}	tʰɐw^{2} tʃi^{3}	li^{2}	–	tʃʰi^{5} tʃi^{3}
NN_p	nam^{3}	taw^{2}	ɬa^{1} luj^{2}	ɬɐj^{6} tʃi^{3}	–
BS_p	lam^{4} ku^{3}	tɐw^{2} tʃi^{3}	ɬa^{1} luj^{2}	ʃɐj^{6} tʃi^{3}	–
BY	kɐm^{3} lam^{3}	tœw^{2} tʃəj^{3}	ləj^{2}	–	ʃɐj^{4} tʃəj^{3}
MS	kəm^{3} lam^{4}	təw^{2} ɬi^{3}	li^{2}	–	ʃi^{4} ɬi^{3}
YJ	lam^{3} tʃij^{3}	tʰɔw^{2} tʃij^{3}	lij^{2}	ʃi^{6} tʃij^{3}	–
WC	ka^{1} lam^{-35}	–	–	–	–

Table C.33

Gloss	pine	plant trees	pomelo	radish	sesame
Index	166	167	168	169	170
Orth	松樹	種樹	–	蘿蔔	芝麻
CY	*dzuŋ^{2} ʒy^{6}	tʃuŋ^{5} ʒy^{6}	*p.luk^{7}	*lɔ^{2} bak^{8}	*tʃi^{1} ma^{2}
YS	tsoŋ^{2} tsy^{6}	tsoŋ^{5} tsy^{6}	pɐt^{-33} lok^{8}	lɔ^{2} pak^{8}	tsi^{1} ma^{2}
LX	tsɔŋ^{2} tʃy^{6}	tʃɔŋ$^{-22}$ tʃy^{6}	pan^{3} lɔk^{7}	lɔ$^{-22}$ pak^{8}	ma^{-22} tsɐj^{3}
HJ	tsuŋ^{2} tsɔw^{6}	tsuŋ^{5} tsɔw^{6}	(kʷɔ^{3} jaw^{5})	lɔ^{2} pak^{8} taw^{2}	tsi^{1} ma^{2}
FK	tsɔŋ^{2} tsy^{6}	tsɔŋ^{5} tsy^{6}	(taj^{6} kam^{1})	luə pak$^{8-35}$	ma^{2} tsɐj^{2}
GZ	tsʰuŋ^{2} sy^{6}	tsuŋ^{5} sy^{6}	luk^{7} jɐw^{3}	lɔ^{2} pak^{8}	tsi^{1} ma^{2}
NH	tsʰuŋ^{2} tsʰy^{6}	tsuŋ^{5} tsʰy^{6}	pɔ^{3} lɔk^{7}	lɔ^{2} pak^{8}	tsi^{1} ma^{2}
DG	tsʰuŋ^{2} sy^{5}	tsuŋ^{5} sy^{5}	vɔ^{2} ŋuk^{7}	ŋɔ^{2} pɛk^{8}	zɐw^{2} ma$^{2-55}$
FG	tsʰuŋ^{2} sy^{6}	tsuŋ^{1} sy^{6}	pɐj^{1} luk^{8}	lɔ^{2} pak^{8}	tsi^{1} ma^{2}
YF	tsʰuŋ^{2} sy^{2}	tsuŋ^{5} sy^{2}	pɐn^{1} luk^{7}	lɔ^{2} pak^{8}	jɐw^{2} ma^{2}
XH	tsʰuŋ^{2} si^{6}	tsuŋ^{1} si^{6}	pu^{3} luk^{7}	lɔ^{2} pak^{8}	tsi^{1} ma^{2}
TS	tʰuŋ^{2} su^{6}	tsuŋ^{1} su^{6}	pu^{3} luk^{7}	lɔ^{2} pak^{8}	tsi^{1} ma^{2}
KP	tʰuŋ^{2} si^{6}	tsuŋ^{1} si^{6}	vu^{3} luk^{7}	lɔ^{2} vak^{8}	tsi^{1} ma^{2}
EP	tsʰuŋ^{2} si^{6}	tsuŋ^{1} si^{6}	pu^{3} luk^{7}	lɔ^{2} pak^{8}	tsi^{1} ma^{2}
BL	tʰuŋ^{2} muk^{8}	tʃuŋ^{5} muk^{8}	pɔ^{1} luk^{7}	!lɔ^{2} puk^{7}	jɐw^{2} ma^{2}
QZ	tʃʰuŋ^{2} muk^{8}	(tʃɔj^{1} ʃi^{6})	puk^{8} tʃi^{3}	!lɔ^{2} puk^{7}	tʃi^{1} ma^{2}
NN_p	tʃuŋ^{2} muk^{10}	(tʃaj^{1} muk^{10})	puk^{8} tʃi^{3}	!la^{2} puk^{9}	tʃi^{1} ma^{2}
BS_p	tʃɔŋ^{2} mɔk^{9}	(tʃaj^{1} ʃuj^{6})	pɔk^{8} tʃi^{3}	!la^{2} puk^{9}	tʃi^{1} ma^{2}
BY	ɬuŋ^{1} ʃuj^{6}	tʃuŋ^{5} ʃuj^{6}	(jəw^{6} tʃəj^{3})	!lœ^{2} pʰuk^{7}	tʃi^{1} ma^{2}
MS	ɬuŋ^{1} muk^{8}	tʃuŋ^{5} tʃi^{6}	lɔk^{8} pɐw^{6}	!lɔ^{2} puk^{7}	tʃi^{1} ma^{2}
YJ	tʃʰuŋ ʃi^{6}	–	puk^{8} tʃij^{3}	lɔ^{2} pɐk^{8}	tʃi^{1} ma^{2}
WC	–	tʃuŋ^{5} ʃi^{6}	ka^{1} pʰuk^{8}	–	ma^{2} ti^{3}

Table C.34

Gloss	sourpeel tangerine	soybean	sugarcane	sweetpeel tangerine	taro
Index	171	172	173	174	175
Orth	橘	黃豆	蔗	柑	芋頭
CY	*kʷit⁷	*wɔŋ² dɐw⁶	*tʃia⁵	*kɔm¹	*wu⁶ dɐw²
YS	kɐt⁷ tsaj³	wɔŋ⁻³³ taw⁶	tse⁵	–	wu⁴ taw²
LX	–	wɔŋ² taw⁶	kan⁻²² tʃɛ⁴	kan⁻²² tsɐj³	wu⁻²² taw²
HJ	kɐt⁷	wɔŋ² taw⁶	tsɛ⁵	–	wu⁶ taw²
FK	tsɐt⁷	wuŋ² tɔw⁶	tsiə⁵	–	wu⁶ tɐw²⁻³⁵
GZ	kɐt⁷	pak⁸ tɐw⁻³⁵	tsɛ⁵	kɐm¹	wu⁶ tʰɐw²⁻³⁵
NH	kɐt⁷	pak⁸ tɐw⁻³⁵	tsɛ⁵	kum¹	fu⁶ taw²
DG	taj⁵ kɐk⁷	pɛk⁸ taw³	tsø⁵	kaŋ¹⁻⁵⁵	fu⁵ tʰaw²
FG	kɐt⁷ tsɐj³	wɔŋ² tɐw⁻³⁵	tse¹	–	wu⁶ tʰɐw²
YF	kɐt⁷ tsɐj³	wɔŋ² tɐw²	tsɛ⁵	–	wu² tʰɐw²
XH	kæt⁷	vaŋ² tæw⁻²⁴	tsæ¹	kæm¹	vu⁶⁻³⁵
TS	kit⁷	vɔŋ² ew⁻³⁵	tsia⁻³⁵	kam⁻³⁵	vu⁶ hew²
KP	kik⁷	vɔŋ² aw⁶	tsia⁻³⁵	kam⁻³⁵	vu⁶ haw²⁻³⁵
EP	kiat⁷	vuaŋ² tij⁻³⁵	tsia⁻³⁵	kam⁻³⁵	vu⁶ hij²⁻³⁵
BL	kʷɐt⁹ ti³	ɔŋ² tɐw⁶	tʃɛ⁵	kom¹ ti³	wu⁻²² tʰɐw²
QZ	kɐt⁷ tʃi³	wɔŋ² tɐw⁶	kam¹ tʃɛ⁵	pɛn³ kam¹	!ji² tʰɐw²
NN_p	kɐt⁹ tʃi³	huŋ² tɐw⁶	kam¹ tʃi⁵	kam¹	!huj⁶ tɐw²
BS_p	kɐt⁹ tʃi³	huŋ² tɐw⁶	kam¹ tʃi⁵	kam¹ tʃi³	!huj² tɐw²
BY	kət⁷	uŋ² tɐw⁶	kam¹ tʃɛ⁵	piŋ¹ kam¹	!huj⁶ tɐw²
MS	kɐt⁷ ɬi³	hɔŋ² tɐw⁶	kəm¹ tʃɛ⁵	kəm¹	!hi² tɐw²
YJ	kɐt⁷ tʃij³	wɔŋ² tɐw⁶	tʃiɛ⁵	kɐm¹	wu⁶ tʰɐw²
WC	kʷɐt⁷ ti³	pʰaak⁸ tʰɐw⁶	tʃɛ⁵	–	wu⁶

Table C.35

Gloss	towel gourd	wheat awn	willow	1sg	2sg
Index	176	177	178	179	180
Orth	絲瓜	麥穗	柳樹	我	你
CY	*sʐ̩^{1} kʷa^{1}	–	liw^{4b} ʒy^{6}	*ŋɔ^{4a}	*ni^{4a}
YS	si^{1} kʷa^{1}	mɐk^{8} joŋ$^{-35}$	lɐw^{4} tsy^{6}	ŋɔ$^{-24}$	nɐj^{-24}
LX	sɐj^{-22} kʷa^{1}	miak^{-22} tʃʰyn^{5}	lɐw^{4} tʃy^{6}	ŋɔ^{4}	nɐj^{4}
HJ	θɐj^{2} kʷa^{1}	syt^{9} wɔ^{2} θin^{5}	law^{2} tsɔw^{6}	ŋɔ^{4}	nɐj^{4}
FK	sɐj^{1} kʷa^{1}	mek^{9} lœŋ^{1}	lɐw^{4} tsy^{6}	(nɔŋ^{1})	nɐj^{2}
GZ	si^{1} kʷa^{1}	mɐk^{8} sœj^{6}	lɐw^{4} sy^{6}	ŋɔ^{4}	nij^{4}
NH	(siŋ^{5} kʷa^{1})	–	lɐw^{4} tsʰy^{6}	ŋɔ^{4}	nɐj^{4}
DG	sɐj^{1} kʷa^{1}	mɛk^{8} suj^{5}	ŋaw^{4} sy^{5}	ŋɔ^{4}	nɐj^{4}
FG	θy^{1} kʷa^{1}	mɐk^{8} θin^{5}	lɐw^{4} sy^{6}	ŋɔ^{3}	ni^{3}
YF	si^{1} kʷa^{1}	mak^{8} juŋ^{1}	lɐw^{4} sy^{2}	ŋɔ^{4}/(nɐn^{1})	nij^{4}
XH	si^{1} kʷa^{1}	mak^{8} suj^{6}	læw^{4} si^{6}	ŋɔ^{1}	nij^{1}
TS	(sen^{1} ka^{1})	mak^{8} ɬuj$^{6\text{-}21}$	liw^{4} su^{6}	(ŋɔj^{-31})	ni^{1}
KP	ɬi^{1} ka^{1}	mak^{8} ɬin^{1}	liaw^{4} si^{6}	(ŋɔj^{1})	nij^{1}
EP	su^{1} ka^{1}	mak^{8} fuj^{6}	lij^{6} si^{6}	ŋua^{6}	ni^{1}
BL	ɬi^{1} kʷa^{1}	mɛk^{-22} ɬin^{5}	ɛŋ^{2} lɐw^{4}	ŋɔ^{4}	ni^{4}
QZ	ɬi^{1} kʷa^{1}	mak^{8} ɬin^{5}	lɐw^{4} muk^{8}	ŋɔ^{4}	ni^{4}
NN_p	ɬi^{1} kʷa^{1}	mɛk^{10} tʃʰun^{5}	lɐw^{4} ɬuj^{6}	ŋa^{4}	nuj^{4}
BS_p	ʃuj^{-55} kʷa^{1}	mɛk^{4} tʃʰun^{5}	lɐw^{4} ʃuj^{6}	ŋa^{4}	ɲuj^{4}
BY	ɬəj^{1} kʷa^{1}	mak^{10} ɬuj^{6}	ləw^{4} ʃuj^{6}	ŋœ^{4}	nəj^{4}
MS	ɬi^{1} kua^{1}	mɛk^{-22} tʃʰun^{5}	lɐw^{4} tʃi^{6}	ŋ̍^{4}	ni^{4}
YJ	ɬij^{1} ka^{1}	–	lɐw^{3} ʃi^{6}	ŋɔ^{3}	nij^{3}
WC	–	–	–	ŋɔ^{4}	ni^{4}

Table C.36

Gloss	3sg	cold (of water)	dirty	dry	insipid
Index	181	182	183	184	185
Orth	佢	凍	邋遢	乾	淡
CY	*gy^{4a}	*tuŋ^{5}	*lat^{8} tʰat^{7}	*kɔn^{1}	*dam^{4a}
YS	kœj^{-24}	toŋ$^{-33}$	lap^{-33} tat^{7}	kon$^{1-52}$	tam^{-24}
LX	kœj^{4}	(liaŋ^{4})	lat^{8} tat^{9}	kɔn^{1}	tan^{4}
HJ	kʰœj^{4}	tuŋ^{5}	lat^{8} tʰat^{9}	kɔn^{1}	tam^{4}
FK	ky^{2}	tɔŋ$^{-33}$	(uə$^{-33}$ tsɔ^{1})	kɔn^{1}	tam^{4}
GZ	kʰœj^{4}	tuŋ^{5}	lat^{8} tʰat^{9}	kɔn$^{1-53}$	tʰam^{4}
NH	kʰœj^{4}	ʈuŋ^{5}	(wu^{1} tɔ^{1})	kɔŋ^{1}	tʰam^{4}
DG	kʰuj^{4}	tuŋ^{5}	(ŋaj^{5} sɔk^{7})	kun^{1}	tʰaŋ^{4}
FG	kʰøj^{3}	tuŋ^{1}	lat^{8} tʰat^{9}	kɔn^{1}	tʰam^{4}
YF	kʰœj^{4}	tuŋ^{5}	lak^{8} tʰak^{9}	kun^{1}	tʰam^{4}
XH	kʰuj^{1}	tuŋ^{1}	(la^{2} tsa^{3})	kun^{1}	ham^{4}
TS	kʰuj^{1}	uŋ^{1}	lat^{8} tʰat^{9}	kɔn^{1}	ham^{1}
KP	kʰuj^{1}	uŋ^{1}	lat^{8} tʰat^{9}	kɔn^{1}	ham^{1}
EP	kʰuj^{1}	tuŋ^{1}	lat^{6} hat^{9}	kuan^{1}	ham^{1}
BL	kʰy^{4}	tuŋ^{5}	lat^{8} tʰat^{9}	kɔn^{1}	tʰam^{3}
QZ	kʰi^{4}	(jɐn^{1})	lat^{8} tʰat^{9}	kɔn^{1}	tʰam^{4}
NN_p	kuj^{2}	(jən^{-55})	(uk^{9} tʃʰuk^{9})	kan^{1}	tam^{4}
BS_p	kuj^{2}	(jən^{1})	(uk^{9} tʃʰuk^{9})	kan^{1}	tam^{4}
BY	na^{6}	(laŋ^{4})	lat^{8} tʰat^{9}	kœn^{1}	tam^{4}
MS	ki^{2}	tuŋ^{5}	(u^{5})	law^{1}	tam^{4}
YJ	kʰij^{2}	–	(wɔ^{5})	kɔn^{1}	tʰam^{3}
WC	ki^{4}	–	–	–	–

Table C.37

Gloss	late	not viscous	pretty	arm	armpit
Index	186	187	188	189	190
Orth	遲	稀	靚	手臂	胳肋底
CY	*dʒi^{2}	*xi^{1}	*ʔliaŋ^{5}	*ʃiw^{3} pi[^{5}]	*k.lak^{7} tɐj^{3}
YS	tsi^{2}	(weŋ$^{-22}$)	leŋ$^{-22}$	sɐw^{3} pɐi^{6}	(kap^{9} ha^{6})
LX	tʃi^{2}	(waŋ^{5})	liŋ^{5} muj^{5}	ʃɐw^{3} k^{w}a^{1}	kak^{8} lak^{7} taj^{3}
HJ	tsi^{2}	(wɛn^{1})	lɛŋ^{5}	θaw^{3} pɐj^{5}	k^{h}ak^{9} lak^{7} ŋam^{5}
FK	tsi^{2}	(ts^{h}iŋ^{1})	liŋ^{5}	sɔw^{3} pej^{5}	lɐk^{7} kik^{7} a$^{4-35}$
GZ	ts^{h}i^{2}	hij$^{1-53}$	lɛŋ^{5}	sɐw^{3} pij^{5}	kak^{9} lak^{7} tɐj^{3}
NH	(aŋ^{5})	hi^{1}	lɛŋ^{5}	–	kɐp^{-33} lɐp^{7} ʈɐj^{3}
DG	ts^{h}i^{2}	(ts^{h}ɛŋ^{1})	ŋøŋ^{5}	saw^{3} pɐj^{5}	(kɔ^{9} ts^{h}ik^{7} ha^{5})
FG	ts^{h}i^{2}	hi^{1}	lɛŋ^{1}	sɐw^{3} pɐj^{4}	(lɔk^{9} wɐt^{7})
YF	ts^{h}i^{2}	hij^{1} liw^{1} liw^{1}	lɛŋ^{5}	sɐw^{3} pij^{5}	ka^{1} lak^{7}
XH	ts^{h}i^{2}	hij^{1}	læŋ^{1}	sæw^{3} pij^{6}	kak^{-33} fat^{7} tæj^{3}
TS	ts^{h}i^{2}	hi^{1}	liaŋ^{1}	siw^{3} pi^{6}	(kap^{-33} ^{n}gap^{7})
KP	ts^{h}i^{2}	hij^{1}	liaŋ^{1}	siw^{3} pij^{6}	(kap^{-33} ^{n}gap^{7})
EP	ts^{h}i^{2}	hi^{1}	liaŋ^{1}	siw^{3} pi^{6}	(kap^{-33} ^{n}gap^{7})
BL	tʃ^{h}i^{2}	hi^{1}	lɛŋ^{1}	ʃɐw^{3} pi^{5}	lɐk^{8} tʃɐk^{7} kɛ^{2}
QZ	tʃ^{h}i^{2}	hi^{1}	liɛŋ^{5}	ʃɐw^{3} pi^{5}	lɐk^{7} li^{6}
NN_p	tʃi^{2}	huj^{1}	lɛŋ^{5}	ɬɐw^{3} puj^{5}	(ɐk^{9} ɛ^{55})
BS_p	tʃi^{2}	huj^{1}	lɛŋ^{5}	ʃɐw^{3} puj^{5}	(ak^{9} ɛ$^{-55}$)
BY	tʃəj^{2}	həj^{1}	lɛŋ^{5}	ʃəw^{3} pɐj^{5}	kɐk^{8} lɐk^{7}
MS	tʃi^{2}	hi^{1}	liəŋ^{5}	(ʃɐw^{3} kiŋ^{3})	(ha^{-22} ja^{5})
YJ	tʃ^{h}i^{5}	hij^{1}	liŋ^{5}	ʃɐw^{3} pij^{5}	–
WC	–	hi^{1}	–	ʃɐw^{3} bi^{-35}	ka^{1} lɐt^{7}

Table C.38

Gloss	belly	body	bone	braid (hair)	breasts
Index	191	192	193	194	195
Orth	肚腩	身	骨	辮	☒
CY	*du[^{4b}] nam^{4}	*ʃin^{1}	*kut^{7}	*p[ɛ]n^{1}	*ʔnɛn^{1}
YS	tɐw^{4} nam^{3}	sɐn^{1}	kʷɐt^{7}	pin^{3}	nen^{-22}
LX	tɐw^{6} nan^{4}	ʃɐn^{1}	kʷɐt^{7}	pin^{6} tsɐj^{3}	nɛn^{4}
HJ	tɔw^{3} nam^{4}	θɐn^{1}	kʷɐt^{7}	pin^{1}	nin^{1}
FK	(tɔ^{4})	sɐn^{1}	kʷɐt^{7}	pen^{4}	nin^{1}
GZ	tʰuw^{4} nam^{4}	sɐn^{1}	kʷɐt^{7}	pin^{1}	nin$^{1\text{-}55}$
NH	tʰuw^{4} nam^{4}	sɐn^{1}	kʷɐt^{7}	pin^{1}	(na^{4})
DG	si^{3} tɔw^{3} naŋ^{4}	sɐn^{1}	kʷɐk^{7}	pin^{1}	nin^{-55}
FG	tu^{3} nam^{6}	sɐn^{1}	kʷɐt^{7}	pin^{1}	nin^{1}
YF	(tuw^{4})	sɐn^{1}	kʷɐt^{7}	pin^{1}	nin^{1}
XH	(tæw^{3} nam^{6})	sæn^{1}	kʷæt^{7}	pin^{1}	nin^{3}
TS	(ɬiaw^{3} u^{3})	sin^{1}	kut^{7}	pin$^{1\text{-}35}$	nen^{-35}
KP	(ɬiw^{3} u^{3} haw^{2})	sin^{1}	kut^{7}	viŋ^{3}	nin^{-35}
EP	(tu^{3} nam^{6})	si[e]n^{1}	kut^{7}	pian$^{1\text{-}35}$	ni(e)n^{-35}
BL	ɬiw^{3} tʰu^{4}	ʃɐn^{1} pʰuj^{1}	kʷɐt^{7} tʰɐw^{2}	!tʰɐw^{2} pʰɛn^{4}	nɛn^{2}
QZ	tʰu^{4} pɐw^{1}	ʃɐn^{1}	kʷɐt^{7} tʰɐw^{2}	ma^{1} pin^{1}	(mɛ^{2})
NN_p	tɔ^{4} nam^{4}	ɬɐn^{1}	kʷɐt^{9} tɐw^{2}	pɛn^{4} tʃi^{3}	(mɛ$^{-21}$)
BS_p	ɬɐj^{5} tɔ^{4}	ʃɐn^{1} tʃi^{3}	kʷɐt^{9} tɐw^{2}	!tɐw^{2} pɛn^{4}	(pɛ^{5})
BY	ɬiw^{3} təw^{4}	ʃən^{1} tʰɐj^{3}	kʷɐt^{7} tɐw^{2}	!tɐw^{2} pin^{6}	(mɛ$^{-45}$)
MS	tu^{4}	ʃən^{1}	kʷɐt^{7} tɐw^{2}	!tɐw^{2} pin^{4}	niŋ^{2}
YJ	tʰɔw^{3} nam^{3}	ʃɐn^{1} tʰɐj^{3}	kɐt^{7}	!pin^{6} tʃij^{3}	niɛn^{5}
WC	tʰu^{4}	ʃɐn^{1}	–	–	nɛŋ^{1}

Table C.39

Gloss	cavity (teeth)	chest	chin	copulate (vulgar)	cowlick (hair)
Index	196	197	198	199	200
Orth	爛牙	胸	下巴	屌	旋
CY	*lan^{6} ŋa^{2}	*xuŋ^{1}	*ɣa^{6} ba^{2}	*t[ia]w^{3}	*dziɔn^{6}
YS	(tsy^{5} ŋa^{2})	hoŋ^{1} haw^{3}	ha^{6} pa^{2}	tew^{3}	taw^{2} fat^{9} tsyn^{6}
LX	lan^{6} ŋa^{2}	hɔŋ^{1}	ha^{6} pa^{2}	tiw^{4}	–
HJ	(tsuŋ^{2} tsy^{5} ŋa^{2})	huŋ^{1} pu^{6}	ha^{6} pa^{2}	tiw^{3}	tsyn^{6}
FK	(tsɔŋ^{2} tsy^{5} ŋa^{2})	sɐm^{1} hɔŋ^{1}	a^{6} pa^{2}	tiw^{3}	tsyn^{6}
GZ	lan^{6} ŋa^{2}	huŋ^{1} puw^{6}	ha^{6} p^{h}a^{2}	tiw^{3}	tsyn^{6}
NH	laŋ^{6} ŋa^{2}	huŋ^{1}	ha^{6} pa^{2}	tiw^{3}	tsyn^{6}
DG	ŋɛŋ^{5} ŋa^{2}	(sɐm^{1} haw^{3})	ha^{6} p^{h}a^{2}	tiw^{1}	nɔw^{4} hɔ^{9} tsøn^{3}
FG	lan^{6} ŋa^{2}	huŋ^{1} hɐw^{3}	ha^{6} p^{h}a^{2}	tiw^{3}	tsyn^{1}
YF	laŋ^{2} ŋa^{2}	huŋ^{1}	ha^{2} p^{h}a^{2}	t^{h}iw^{3}	tsyn$^{6\text{-}35}$
XH	lan^{6} ŋa^{2}	huŋ^{1} hæw^{3}	ha^{6} p^{h}a^{2}	tiw^{3}	tsin^{6}
TS	lan^{6} ŋa^{2}	huŋ^{1} hew^{3}	ha^{6} p^{h}a^{2}	iw^{5}	tun$^{6\text{-}21}$
KP	lan^{6} ŋa^{2}	huŋ^{1} haw^{3}	ha^{6} ha^{2}	ew^{5}	(iŋ^{3})
EP	lan^{6} ŋa^{2}	huŋ^{1} hij^{3}	ha^{6} p^{h}a^{2}	tiw^{1}	ts^{h}ien^{6}
BL	(tʃ^{h}uŋ^{2} ŋa^{2})	huŋ^{1} p^{h}u^{2}	ha^{6} p^{h}a^{2}	tiw^{3}	(t^{h}ɐw^{2} lo^{2})
QZ	(tʃ^{h}uŋ^{2} ŋa^{2})	huŋ^{1} p^{h}u^{2}	ha^{6} p^{h}a^{2}	tiw^{3}	t^{h}ɐw^{2} tʃin^{6}
NN_p	(tʃuŋ^{2} ɲa^{2})	huŋ^{1} pu^{2}	ja^{6} pa^{2}	tiw^{3}	tɐw^{2} fat^{9} tʃin^{6}
BS_p	(tʃɔŋ^{2} ɲa^{2})	hɔŋ^{1} taŋ^{2}	ja^{6} pa^{2}	tiw^{3}	tʃin^{3} tʃin^{6} luj^{2}
BY	(tʃuŋ^{2} ɲa^{2})	huŋ^{1} tœŋ^{2}	ja^{4} pa^{2}	tiw^{3}	tɐw^{2} tʃun^{3} tʃun^{6}
MS	(ŋa^{2} tʃuŋ^{2})	(ɬɐm^{1} hɐw^{3})	ja^{6} pa^{2}	tiw^{3}	tɐw^{2} tʃun^{6}
YJ	–	(ɬɐm^{1} hɐw^{3})	ha^{6} pa^{1}	tiw^{3}	tʃun^{6}
WC	–	(ɬɐm^{1} hɐw^{3})	ha^{6} dit^{9}	–	–

Table C.40

Gloss	deaf	ears	eye	eyebrow	face
Index	201	202	203	204	205
Orth	耳背	耳	眼	眉	面
CY	*ʔɲi³ buɔj⁶	*ʔɲi³	*ʔŋan³	*mi²	*mian⁶
YS	ji⁴ puj⁶	ji⁴ tsɐj³	ŋen⁴	ŋen⁴-⁴⁴ mɐj²	min⁶
LX	ɲi⁴ puj⁵	ɲi⁴	ŋɛn⁴	mɐj² mɔ²	min⁶
HJ	(ɲi⁴ pʰɐŋ¹ θɐŋ¹)	ɲi⁴ tsɐj³	ŋɛn⁴	ŋɛn⁴ mɐj²	min⁶
FK	(ɲ(ʲ)i⁴ luŋ¹)	ɲ[ʲ]i⁴	ŋan⁴	ŋan⁴ mɐj²	min⁶
GZ	ji⁴ puj⁶	ji⁴ tsɐj³	ŋan⁴	ŋan⁴ mij²	min⁶
NH	ji⁴ py⁵	ji⁴ tsɐj³	ŋɛŋ⁴	ŋɛŋ⁴ mɐj²	min⁶
DG	ŋɐj⁴ kɐk⁷ puj⁵	ŋɐj⁴ kɐk⁷	ŋɛŋ⁴	ŋɛŋ⁴ mɐj² mɔw²	min⁵
FG	ji⁴ puj¹	ji⁴	ŋan³	ŋan³ mɐj²	min⁶-³⁵
YF	ji⁴ pu⁵	ji⁴ tsɐj³	ŋaŋ⁴	ŋaŋ⁴ mij²	min²
XH	ŋi³ puj¹	ŋi³ tsæj³	ŋan³	mij² mu²	min⁶
TS	ŋ[e]i³ pɔj¹	ŋ[e]i¹ tɔj³	ŋan³	m[e]i² maw²	men⁶
KP	ŋij³ vɔj¹	ŋij³ tɔj³	ŋan³	mij² mɔ²	min⁶
EP	ŋi³ puaj¹	ŋi³ tsaj³	ŋan³	mi² maw²	mian⁶
BL	ɲi⁴ puj⁶	ɲi⁴ to³	ŋan⁴	mi² mɐw²	min⁶
QZ	ji⁴ puj⁶	ji⁴ tɔ³	ŋan⁴	ŋan⁴ mi² maw²	min⁶
NN_p	ɲi⁴ pɔj⁶	ɲi⁴ tu³	ɲan⁴	muj² maw²	min⁶
BS_p	ɲi⁴ pɔj⁶	ɲi⁴ tu³	ɲan⁴	muj² maw²	min⁶
BY	ɲi⁴ puj⁶	ɲi⁴ təw³	ɲan⁴ tʃiŋ¹	mɐj² mœw²	min⁶
MS	ŋi⁴ puj⁶	ŋi⁴ tɔ³	ŋan⁴	ŋan⁴ mi² məw²	min⁶
YJ	ji³ puj⁶	ji³	ŋan³	ŋan³ mij²	min⁶
WC	–	ɲi⁴	ŋan⁴	!mi²	min⁶

Table C.41

Gloss	facial hair	finger	fingernail	forehead	gizzard
Index	206	207	208	209	210
Orth	鬚	手指	甲	額	腎
CY	*su^{1}	*ʃiw^{3} tʃi^{3}	*kap^{7}	*ŋak^{8}	*gin$^{4a/b}$
YS	vu^{2} sɐw^{1}	sɐw^{3} tsi^{3}	tsi^{3} kap^{9}	ŋek^{8} taw^{2}	kɐn^{5}
LX	!sœj^{1}	ʃɐw^{5} tʃi^{3}	ʃɐw^{5} tʃi^{3} kɛt^{9}	ŋiak^{8} taw^{2}	kɐn^{4}
HJ	wu^{2} θɔw^{1}	θaw^{3} tsi^{3}	tsi^{3} kɛt^{9}	ŋɛk^{8} taw^{2}	kɐn^{4}
FK	!wu^{2} sœj^{1}	sɔw^{3} tsi^{3} tɐw^{2}	sɔw^{3} tsi^{3} kap^{9}	ŋak^{8} tɐw^{2}	tsɐn^{4}
GZ	suw^{1}	sɐw^{3} tsi^{3}	sɐw^{3} tsi^{3} kap^{9}	ŋak^{8} tʰɐw^{2}	(sɐn^{4})
NH	wu^{2} suw^{1}	sɐw^{3} tsi^{3}	sɐw^{3} kɛp^{9}	ŋak^{8} taw^{2}	(tsʰɐn^{4})
DG	fu^{2} sɔw^{1}	saw^{3} tsi^{3}	saw^{3} tsi^{3} ka^{5}	ŋɛk^{8}	kʰɐn^{4}
FG	!wu^{2} søj^{1}	sɐw^{3} tsi^{3}	tsi^{3} kap^{9}	ŋak^{8} tʰɐw^{2}	kʰɐn^{4}
YF	wu^{2} suw^{1}	sɐw^{3} tsi^{3}	tsi^{3} kap^{9}	tsʰin^{2} ŋak^{8}	kʰɐn^{4}
XH	sæw^{1}	sæw^{3} tsi^{3}	sæw^{3} kap$^{9\text{-}11}$	ŋak^{8} hæw^{2}	kʰæn^{1}
TS	ɬu^{1}	siw^{3} tsi^{3}	siw^{3} kap$^{9\text{-}11}$	m[e]i^{2} ŋak^{8}	kʰin^{4}
KP	ɬu^{1}	siw^{3} tsi^{3}	siw^{3} kap$^{9\text{-}11}$	mij^{2} ŋak^{8}	kʰiŋ^{4}
EP	su^{1}	siw^{3} tsi^{3}	siw^{3} tsi^{3} kap$^{9\text{-}11}$	mi^{2} ŋak^{8}	kʰian^{6}
BL	wu^{2} ɬu^{1}	ʃɐw^{3} ɲi^{2}	ʃɐw^{3} kap^{9}	ŋɛk$^{\text{-}22}$ tʰɐw^{2}	(ʃɐn^{4})
QZ	fu^{2} ɬu^{1}	ʃɐw^{3} tʃi^{3}	tʃi^{3} kap^{9}	ŋak^{8} tʰɐw^{2}	(ʃɐn^{4})
NN_p	(hu^{2} tʃi^{3})	ɬɐw^{3} tʃi^{3}	tʃi^{3} kɛp^{9} puŋ^{2}	ŋɛk^{10} tɐw^{2}	ɬɐn$^{\text{-}24}$
BS_p	!hu^{2} ɬə^{1}	ʃɐw^{3} tʃi^{3}	tʃi^{3} kap^{9}	ŋɛk^{4} tɐw^{2}	ʃɐn^{4}
BY	!hu^{2} ɬuj^{1}	ʃəw^{3} tʃi^{3}	tʃi^{3} tʃap^{9}	ŋak^{10} tɐw^{2}	(ʃɐn^{4})
MS	!hu^{2} ɬi^{1}	ʃɐw^{3} ŋɛ^{1}	tʃi^{3} kap^{7}	tɐw^{2}	(tʃɐn^{4})
YJ	wu^{2} ɬɔw^{1}	ʃɐw^{3} tʃi^{3}	ʃɐw^{3} kap^{9}	ŋak^{8} tʰɐw^{2}	–
WC	–	ʃɐw^{3} ɲi^{2}	–	ŋak^{8} tʰɔŋ^{2}	–

Table C.42

Gloss	grasp in armpit	hair (head)	hand	head	heel
Index	211	212	213	214	215
Orth	擸	頭毛	手	頭	腳踁
CY	tʃ[ia]t^{7}	*dɐw^{2} mɔw^{2}	*ʃiw^{3}	*dɐw^{2}	*kiɔk^{7} tʃaŋ^{1}
YS	(tsi^{3})	(taw^{2} fat^{9})	sɐw^{3}	taw^{2}	kœk^{9} tsɐŋ^{1}
LX	tsit^{7}	(taw^{2} fat^{9})	ʃɐw^{3}	taw^{2}	kyk^{8} kɐn^{1}
HJ	(θi^{3})	(taw^{2} fat^{9})	θaw^{3}	taw^{2}	(kœk^{9} ŋɛŋ^{1})
FK	–	(tɐw^{2} fat^{9})	sɔw^{3}	tɐw^{2}	(kik^{7} tok^{7})
GZ	tsit^{7}	tʰɐw^{2} muw^{2}	sɐw^{3}	tʰɐw^{2}	kœk^{9} tsaŋ^{1}
NH	tsit^{7}	(taw^{2} pʰɐt^{9})	sɐw^{3}	taw^{2}	kœk^{9} tsaŋ^{1}
DG	tsit^{7}	tʰaw^{2} mɔw^{2}	saw^{3}	(nɔw^{4} hɔ^{5})	kø^{9} tsaŋ^{1}
FG	tsit^{7}	(tʰɐw^{2} fat^{9})	sɐw^{3}	tʰɐw^{2}	køk^{9} tsaŋ$^{1\text{-}35}$
YF	(kit^{9})	(tʰɐw^{2} fak^{9})	sɐw^{3}	tʰɐw^{2}	kœk^{9} tsɐŋ^{1}
XH	tsit^{7}	hæw^{2} mu^{2}	sæw^{3}	hæw^{2}	kiɔk tsaŋ^{1}
TS	tsit^{7}	hew^{2} maw^{2}	siw^{3}	hew^{2}	kiak^{9} tsaŋ^{1}
KP	(kit^{7})	haw^{2} mɔ^{2}	siw^{3}	haw^{2}	kiak^{9} haw^{6} tsaŋ^{1}
EP	(ki[e]t^{7})	hij^{2} maw^{2}	siw^{3}	hij^{2}	(kiɔk^{9} hij^{2})
BL	tʃɛ^{1}	tʰɐw^{2} mɐw^{2}	ʃɐw^{3}	tʰɐw^{2}	(kɛk^{-33} tuk^{7})
QZ	tʃɛt^{1}	tʰɐw^{2} maw^{2}	ʃɐw^{3}	tʰɐw^{2}	(kɛk^{9} tuk^{7})
NN_p	(jɐk^{7})	tɐw^{2} maw^{2}	ɬɐw^{3}	tɐw^{2}	(kɛk^{9} tuk^{9})
BS_p	(ni^{4} nɛm^{-35})	tɐw^{2} maw^{2}	ʃɐw^{3}	tɐw^{2}	(kɛk^{9} kɐn^{1})
BY	(mɐj^{5})	tɐw^{2} mœw^{2}	ʃəw^{3}	tɐw^{2}	(kɛk^{9} kɐn^{1})
MS	(lɔ^{1} lɛm^{1})	(tɐw^{2} fat^{7})	ʃɐw^{3}	tɐw^{2}	(kiək^{7} tuk^{7})
YJ	–	tʰɐw^{2} mɔw^{2}	ʃɐw^{3}	tʰɐw^{2}	kiɛk^{9} tʃaŋ^{1}
WC	–	(tʰɐw^{2} fat^{9})	ʃɐw^{3}	tʰɐw^{2} kwɐt^{7}	kɛk^{9} tʃaŋ^{1}

Table C.43

Gloss	knee	left	leg	lips	mouth
Index	216	217	218	219	220
Orth	膝頭哥	左	腳	脣	嘴
CY	*[s]i[t]^{7} dɐw^{2} kɔ^{1}	*tsɔ^{3}	*kiɔk^{7}	*ʒyn^{2}	*tsuj^{3}
YS	ts^{h}ɐt^{7} taw^{2} ku^{3}	tsɔ^{3}	kœk^{9}	tsoj^{3} sɐn^{2}	tsoj^{3}
LX	ts^{h}ɐt^{8} taw^{6} kœj^{5}	tsɔ^{3}	kiak^{7}	haw^{3} tʃɐn^{2}	tsœj^{3}
HJ	θɐt^{7} taw^{2} ku^{3}	tsɔ^{3}	kœk^{9}	tsɔj^{3} tsɐn^{2}	tsɔj^{3}
FK	sɐt^{7} tɐw^{2} ku^{1}	tsuə^{3}	kik^{7}	tsœ^{3} tsɐn^{2}	tsœ^{3}
GZ	sɐt^{7} t^{h}ɐw^{2} kɔ^{1}	tsɔ^{3}	kœk^{9}	tsœj^{3} sœn^{2}	tsœj^{3}
NH	sɐt^{7} taw^{2} kɔ^{1}	tɔ^{3}	kœk^{9}	hɐw^{3} tsœn^{2}	tɔ^{3}
DG	sɐk^{7} t^{h}aw^{2} kɔ^{1}	tsɔ^{3}	kø^{9}	tsuj^{3} sɐŋ^{2}	tsuj^{3}
FG	!ts^{h}ɐt^{7} t^{h}ɐw^{2} kɔ^{1}	tsɔ^{3}	køk^{9}	tsøj^{3} sɐn^{2}	tsøj^{3}
YF	!ts^{h}ɐt^{7} t^{h}ɐw^{2} kɔ^{1}	tsɔ^{3}	kœk^{9}	tsu^{3} sɐn^{2}	tsu^{3}
XH	!sæp^{7} hæw^{2} kɔ^{1}	tsɔ^{3}	kiɔk	tsuj^{3} sæn^{2}	tsuj^{3}
TS	!ɬip^{7} hew^{2} kɔ^{1}	tɔ^{3}	kiak^{9}	hew^{3} sun^{2}	tuj^{3}
KP	!ɬ(e)p^{7} haw^{2} ku^{3}	tu^{3}	kiak^{9}	haw^{3} sun^{2}	tuj^{3}
EP	!siap^{7} hij^{2} kua^{3}	tsua^{3}	kiɔk^{9}	hij^{3} sun^{2}	tsuj^{3}
BL	ɬɐt^{7} t^{h}ɐw^{2}	tɔ^{3}	kɛk^{-33}	tuj^{3} ʃɐn^{2}	tuj^{3}
QZ	ɬɐt^{7} t^{h}ɐw^{2}	tʃɔ^{3}	kɛk^{9}	tʃuj^{3} ʃɐn^{2}	tʃuj^{3}
NN_p	ɬɐt^{9} tɐw^{2} kuŋ^{1}	tʃa^{5}	kɛk^{9}	hɐw^{3} ɬɐn^{2}	tʃuj^{3}
BS_p	ɬɐt^{9} tɐw^{2}	tʃa^{5}	kɛk^{9}	hɐw^{3} ʃən^{2}	(hɐw^{3})
BY	ɬət^{7} tɐw^{2}	tʃœ^{5}	kɛk^{9}	tʃœ^{3} ʃən^{2}	tʃœ^{3}
MS	(pɔ^{1} lɔ^{2} kuaj^{5})	ɬɔ^{3}	kiək^{7}	hɐw^{3} tʃən^{2}	(hɐw^{3})
YJ	ɬɐt^{7} t^{h}ɐw^{2}	tʃɔ^{3}	(t^{h}uj^{3})	tʃuj^{3} ʃɐn^{2}	tʃuj^{3}
WC	ɬɐk^{7} t^{h}ɐw^{2}	tɔ^{3}	kɛk^{9}	–	–

Table C.44

Gloss	navel	neck	nose	nostril	pupil
Index	221	222	223	224	225
Orth	臍	頸	鼻	鼻哥窿	眼☒
CY	*dzi^{2}	*kiaŋ^{3}	*bi^{6}	*bi^{6} kɔ^{1} ʔluŋ^{1}	*ʔŋan^{3} wut^{8}
YS	tɐw^{4} tsɐj^{2}	kiŋ^{3}	pɐj^{-22} koŋ$^{1-52}$	pɐj^{6} luŋ^{1}	ŋen$^{4-44}$ wɐt^{8}
LX	tɐw^{6} tsɐj^{2}	kiŋ^{3}	pɐj^{6}	pɐj^{6} lɔŋ^{1}	ŋɛn^{4} tʃy^{1}
HJ	tɔw^{3} tsɐj^{1}	kɛŋ^{3}	pɐj^{6} kɔ^{1}/pɐj^{6} lɐŋ^{5} nɐŋ^{5}	pɐj^{6} kɔ^{1} luŋ^{1}	ŋɛn^{4} wɐt^{8}
FK	tɔ^{4} tsɐj^{2}	kiŋ^{3}	pej^{6} kœ^{1}	pej^{6} kœ^{1} luŋ^{1}	ŋan^{4} wɐt^{8}
GZ	tʰuw^{4} tsʰi^{2}	kɛŋ^{3}	pij^{6} kɔ$^{1-55}$	pij^{6} kɔ^{1} luŋ^{1}	ŋan^{4} wɐt^{8}
NH	tʰuw^{4} ty^{2}	kɛŋ^{3}	pɐj^{6} kœ^{1}	pɐj^{6} kɔ^{1} luŋ^{1}	ŋɛŋ^{4} wɐt^{8}
DG	tɔw^{4} tsʰɐj^{2}	køŋ^{3}	pɐj^{5} kɔ$^{-55}$	pɐj^{5} kɔ^{1} ŋuŋ$^{1-11}$	ŋɛŋ^{4} fɐk^{8}
FG	tu^{3} tsʰɐj^{2}	kiŋ^{3}	pɐj^{6} kɔ^{1}	pɐj^{6} kɔ^{1} luŋ^{1}	ŋan^{3} wɐt^{8}
YF	tuw^{4} tsʰi^{2}	kɛŋ^{3}	pij^{2} kɔ^{1}	pij^{2} kœ^{1} luŋ^{1}	ŋaŋ^{4} wɐt^{8}
XH	tæw^{3} tsʰæj^{2}	kæŋ^{3}	pij^{6} kɔ^{3}	(pij^{6} kʰuŋ^{3})	ŋan^{3} væt^{8}
TS	u^{3} tʰij^{2}	kiaŋ^{3}	pij^{6} kɔ^{3}	(p[e]i^{6} kʰuŋ^{3})	(ŋan^{3} ŋin^{2})
KP	u^{3} tʰaj^{2}	kiaŋ^{3}	vij^{6} ku^{3}	(vij^{6} kʰuŋ^{3})	(ŋan^{3} ŋiŋ^{2})
EP	tu^{3} tsʰi^{2}	kiaŋ^{3}	pi^{6} kua^{3}	pi^{6} fut^{7} luŋ$^{-55}$	(ŋan^{3} ŋian^{2})
BL	tʰu^{4} tʰɐj^{2}	kiŋ^{3}	pi^{6} kuŋ^{1}	pi^{6} kuŋ^{1} luŋ^{1}	ŋan^{4} hɐt^{8}
QZ	tʰu^{4} tʃʰi^{2}	kiɛŋ^{3}	pi^{6} kuŋ^{1}	pi^{6} luŋ^{1}	ŋan^{4} wɐt^{8}
NN_p	tɔ^{4} tʃɐj^{2}	kiŋ^{3}	(pɐt^{8} tʃi^{3})	pɐt^{8} luŋ^{1}	ɲan^{4} wɐt^{8}
BS_p	tɔ^{4} tʃɐj^{2}	kən^{3}	(puət^{8} tʃi^{3})	p(u)ət^{8} lɔŋ^{1}	ɲan^{4} wət^{-24}
BY	təw^{4} tʃɐj^{2}	kiŋ^{3}	(pɐt^{8})	pɐt^{8} luŋ^{1}	ɲan^{4} wɐt^{8}
MS	(tu^{4} pɛ^{1})	kiŋ^{3}	pi^{6}	pi^{6} kuŋ^{1}	ŋan^{4} wɐt^{8}
YJ	tʰɔw^{3} tʃʰij^{2}	kiŋ^{3}	pij^{6}	pij^{6} luŋ^{1}	ŋan^{3} wɐt^{8}
WC	–	kiŋ^{3}	pʰi^{6}	–	ŋan^{4} wɐt^{8}

Table C.45

Gloss	rear/anus	ribs (1)	ribs (2)	right	saliva
Index	226	227	228	229	230
Orth	屎窟	肋骨	–	右	口水
CY	*ʃi^{3} xut^{7}	*lɐk^{8} kut^{7}	*pʰiaŋ^{1}	*jiw^{6}	*xɐw^{3} ʃuj^{3}
YS	si^{3} fɐt^{7} tyn^{2}	lɐk^{7} kʷɐt^{7}	–	jaw^{6}	haw^{3} sy^{3}
LX	ʃi^{3} fɐt^{7}	lak^{7} kʷɐt^{7}	–	jɐw^{6}	haw^{3} ʃy^{3}
HJ	θi^{3} fɐt^{7}	lak^{8} hɛŋ^{1} kʷɐt^{7}	lak^{8} hɛŋ^{1} kʷɐt^{7}	jaw^{6}	haw^{3} θuj^{3}
FK	si$^{3\text{-}33}$ fɐt^{7}	–	–	jɔw^{6}	hɐw^{3} sœ^{3}
GZ	si^{3} fɐt^{7}	lɐk^{8} kʷɐt^{7}	–	jɐw^{6}	hɐw^{3} sœj^{3}
NH	si^{3} wɐt^{7}	–	pʰɛŋ^{1} kʷɐt^{7}	jɐw^{6}	hɐw^{3} sɔ^{3}
DG	si^{3} fɐk^{7}	ŋɐk^{8} kʷɐk^{7}	–	zaw^{5}	haw^{3} suj^{3}
FG	si^{3} fɐt^{7}	lɐk^{8} pʰaŋ$^{1\text{-}22}$ kʷɐt^{7}	lɐk^{8} pʰaŋ$^{1\text{-}22}$ kʷɐt^{7}	jɐw^{6}	hɐw^{3} søj^{3}
YF	si^{3} wɐt^{7}	–	pʰɛŋ^{1} kʷɐt^{7}	jɐw^{2}	hɐw^{3} su^{3}
XH	si^{3} fæt^{7}	lak^{8} kʷæt^{7}	–	zæw^{6}	hæw^{3} suj^{3}
TS	(hun^{2})	lak^{8} kut^{7}	–	ziw^{6}	hew^{3} suj^{3}
KP	(hun^{2})	lak^{8} kut^{7}	–	ziw^{6}	haw^{3} suj^{3}
EP	(hun^{2})	lak^{8} kut^{7}	pʰiaŋ$^{1\text{-}31}$ tsaj^{3} kut^{7}	ziw^{6}	hij^{3} suj^{3}
BL	ʃi^{3} wɐt^{7}	lɐk^{8} ʃak^{9} kʷɐt^{7}	–	jɐw^{6}	(ʃan^{2} ʃuj^{3})
QZ	ʃi^{3} fɐt^{7}	lɐk^{8} kʷɐt^{7}	–	jɐw^{6}	hɐw^{3} ɬuj^{3}
NN_p	(pʰuj^{5} ɬɔt^{9})	lɐk^{10} kʷɐt^{9}	–	jɐw^{6}	(ɬan^{2})
BS_p	(pʰuj^{5} ɲan^{4} tɐn^{2})	lɐk^{9} kʷɐt^{9}	–	jɐw^{6}	(ʃan^{2})
BY	ʃi^{3} wɐt^{7}	lɐk^{8} kʷɐt^{7}	–	jəw^{6}	(ʃan^{2})
MS	ʃi^{3} fɐt^{7}	lɐk^{8} pʰɛŋ^{1} kʷɐt^{7}	lɐk^{8} pʰɛŋ^{1} kʷɐt^{7}	jɐw^{6}	hɐw^{3} ʃuj^{3}
YJ	ʃi^{3} fɐt^{7}	lɐt^{8} kɐt^{7}	–	jɐw^{6}	hɐw^{3} ʃuj^{3}
WC	ʃi^{3} fɐt^{7}	(ʃat^{9} tʃi^{3} kʷɐt^{7})	–	jɐw^{6}	–

Table C.46

Gloss	shoulder	spine	stomach	tear	teeth
Index	231	232	233	234	235
Orth	膊頭	脊骨	肚	眼淚	牙
CY	*pɔk^{7} dɐw^{2}	*tsiak^{7} kut^{7}	*du[^{4b}]	*ʔŋan^{3} luj^{6}	*ŋa^{2}
YS	pɔk^{-33} taw^{2}	tsik^{7} kʷɐt^{7}	tɐw^{4}	ŋen^{4} loj^{5} sy^{3}	ŋa^{2}
LX	pœk^{8} taw^{2}	tsik^{8} kʷɐt^{7}	tɐw^{4}	ŋɛn^{4} lœj^{6}	ŋa^{6} tʃʰi^{3}
HJ	pɔk^{9} taw^{2}	tsɛk^{7} lœŋ^{2} kʷɐt^{7}	tɔw^{3}	ŋɛn^{4} lɔj^{6}	ŋa^{2}
FK	pœk^{-33} tɐw^{2}	iw^{1} kʷɐt^{7}	tɔ^{4}	ŋan^{4} lœ^{6}	ŋa^{2}
GZ	pɔk^{9} tʰɐw^{2}	tsɛk^{9} kʷɐt^{7}	tʰuw^{4}	ŋan^{4} lœj^{6}	ŋa^{2}
NH	pɔk^{9} tʰaw^{2}	tɛk^{9} kʷɐt^{7}	tʰuw^{4}	(ŋɛŋ^{4} sɔ^{3})	ŋa^{2}
DG	(kɛŋ^{1} tʰaw^{2})	(puj^{5} mui^{21} kʷɐk^{7})	si^{3} tɔw^{-35}	ŋɛŋ^{4} ŋuj^{5}	ŋa^{2}
FG	pɔk^{9} tʰɐw^{2}	puj^{1} tsik^{7} kʷɐt^{7}	!tu^{3}	ŋan^{3} løj^{6}	ŋa^{2}
YF	pɔk^{9} tʰɐw^{2}	tsɛk^{9} kʷɐt^{7}	tuw^{4}	ŋaŋ^{4} lu^{2}	ŋa^{2}
XH	(kæn^{1} hæw^{2})	puj^{1} tsæk^{9} kʷæt^{7}	!tæw^{3}	ŋan^{3} luj^{6}	ŋa^{2}
TS	(kan^{1} hew^{2})	pɔj^{1} tiak^{9} kut^{7}	!u^{3}	ŋan^{3} luj^{6}	ŋa^{2}
KP	(kan^{1} haw^{2})	ziaw^{1} tiak^{9} kut^{7}	!u^{3}	ŋan^{3} luj^{6}	ŋa^{2}
EP	(kan^{1} hij^{2})	ziw^{1} tsiak^{9} kut^{7}	!tu^{3}	ŋan^{3} luj^{6}	ŋa^{2}
BL	pɔk^{9} tʰɐw^{2}	(jiw^{1} kʷɐt^{7})	tʰu^{4}	ŋan^{4} luj^{6} ʃuj^{3}	ŋa^{2} tʃʰi^{3}
QZ	pɔk^{9} tʰɐw^{2}	tʃɛk^{9} kʷɐt^{7}	tʰu^{4}	ŋan^{4} luj^{6} ɬuj^{3}	ŋa^{2} tʃʰi^{3}
NN_p	puk^{9} tɐw^{2}	(iw^{1} kʷɐt^{9})	tɔ^{4}	ɲan^{4} luj^{6} ɬuj^{3}	ɲa^{2} tʃʰi^{3}
BS_p	puk^{9} tɐw^{2}	iw^{1} tʃət^{9} kʷɐt^{9}	tɔ^{4}	ɲan^{4} luj^{6} ʃuj^{3}	ɲa^{2} tʃʰi^{3}
BY	puk^{9} tɐw^{2}	tʃik^{7} kʷɐt^{7}	təw^{4}	ɲan^{4} luj^{6} ʃœ^{3}	ɲa^{2} tʃʰi^{3}
MS	puak^{7} tɐw^{2}	(a^{-22} puj^{5} kʷɐt^{7})	tu^{4}	ŋan^{4} luj^{6} ʃuj^{3}	ŋa^{2} tʃʰi^{5}
YJ	pɔk^{9} tʰɐw^{2}	puj^{5} tʃik^{7} kɐt^{7}	tʰɔw^{3}	ŋan^{3} luj^{6}	ŋa^{2}
WC	–	–	tʰu^{4}	–	–

Table C.47

Gloss	thigh (1)	thigh (2)	thumb	tongue	vulva	wing
Index	236	237	238	239	240	241
Orth	大髀	大腿	手指公	脷	閪	翼
CY	*daj^{6} pi^{3}	*daj^{6} tʰuɔj^{3}	*ʃiw^{3} kuŋ^{1}	*li^{6}	*xɐj^{1}	*jik^{8}
YS	taj^{6} pɐj^{3}	–	sɐw^{3} tsi^{3} taw^{2}	lɐj^{-22}	hɐj^{1}	(tsʰi^{5})
LX	–	taj^{6} tʰœj^{3}	ʃɐw^{5} tʃi^{3} kɔŋ^{1}	ʃit^{8} taw^{2}	hɐj^{1}	(tʃʰi^{6} pœŋ^{3})
HJ	–	taj^{6} tʰœj^{3}	θaw^{3} tsi^{3} taw^{2}	lɐj^{6}	hɐj^{1}	(tsʰi^{5})
FK	–	taj^{6} tʰœ^{3}	sɔw^{3} tsi^{3} tɐw^{2}	tsit^{8} lej^{6}	hɐj^{1}	(tsʰi^{5})
GZ	taj^{6} pij^{3}	–	sɐw^{3} tsi^{3} kuŋ^{1}	lij^{6}	hɐj^{1}	jik^{8}
NH	ta^{6} pɐj^{3}	–	sɐw^{3} kuŋ^{1} tʰaw^{2}	lɐj^{6}	hɐj^{1}	hik^{8}
DG	taj^{5} pɐj^{3}	–	saw^{3} tsi^{3} kuŋ^{1}	ŋɐj^{3}	haj^{1}	zik^{8}
FG	taj^{6} pɐj^{3}	–	sɐw^{3} tsi^{3} kuŋ^{1}	lɐj^{6}	hɐj^{1}	ik^{8}
YF	taj^{2} pij^{3}	–	sɐw^{3} tsi^{3} tʰɐw^{2}	lij^{2}	hɐj^{1}	jik^{8}
XH	taj^{6} pij^{3}	–	sæw^{3} tsi^{3} kuŋ^{1}	lij^{6}	hæj^{1}	zik^{8}
TS	aj^{6} pij^{3}	–	siw^{3} kuŋ$^{1\text{-}21}$	li^{-35}	haj^{1}	zet^{8}
KP	aj^{6} vij^{3}	–	siw^{3} kuŋ$^{1\text{-}21}$	lij^{6}	haj^{1}	zik^{-21}
EP	taj^{6} pi^{3}	–	siw^{3} kuŋ$^{1\text{-}21}$	li^{-35}	haj^{1}	zit^{8}
BL	–	taj^{6} tʰuj^{3}	ʃɐw^{3} mɐw^{6} tʰɐw^{2}	li^{6} tʰin^{6}	hɐj^{1}	jik^{8}
QZ	–	taj^{6} tʰuj^{3}	ʃɐw^{3} na^{3} tʰɐw^{2}	li^{6} tʃʰin^{2}	hɐj^{1}	jik^{8}
NN_p	–	taj^{6} tʰɔj^{3}	ɬɐw^{3} tɐw^{2} kuŋ^{1}	luj^{6} tʃin^{2}	hɐj^{1}	hik^{10}
BS_p	–	taj^{6} tʰɔj^{3}	mɐw^{4} tɐw^{2} nɛŋ^{2}	luj^{6} tʃin^{2}	hɐj^{1}	jət^{4} paŋ^{3}
BY	–	taj^{6} tʰuj^{3}	ʃəw^{3} tʃi^{3} tɐw^{2}	ləj^{-21} tʃin^{2}	hɐj^{1}	jik^{8}
MS	(taj^{6} kiək^{7})	–	ʃɐw^{3} na^{4} tɐw^{2}	(ʃit^{8})	hɐj^{1}	hik^{8} tʃʰi^{5}
YJ	taj^{6} pij^{3}	taj^{6} tʰuj^{3}	ʃɐw^{3} ma^{3} kuŋ^{1}	lij^{6}	hɐj^{1}	jik^{8}
WC	–	–	ʃɐw^{3} muk^{7} tʰɐw^{2}	li^{6}	–	jik^{8}

Table C.48