Encoding Syllables in Zhangzhou Southern Min

Yishan Huang
Linguistics Department
The University of Sydney
yishan.huang@sydney.edu.au

Abstract

This study explores how syllables are encoded in Zhangzhou Southern Min, a Sinitic dialect spoken in the southern part of Fujian province in mainland China. A template of C(G)V(X) can be generalized to characterise the internal structure of synchronic Zhangzhou syllables, in which onset (C) and nucleus (V) are obligatory, while glide (G) and coda (X) are optional. Prevocalic glides possess an independent status, while the postvocalic glides are assigned to be one type of syllable coda. Nuclei incorporate oral vowels, nasalised vowels, and syllabic nasals, while codas comprise glides, nasal consonants, and obstruent codas. Attentions are given to discuss the phonetic property of syllable components, as well as the phonological treatment when there exist competing solutions. The description substantially expands our understanding of the phonetics and phonology of the segmental system in this dialect, while it is expected to serve as a model for thorough investigation of how syllable is encoded across Sinitic dialects using modern linguistic theories and analytical approaches. This study also provides important empirical data to the generalization of areal characteristics within the Sino-Tibetan language family, and to the typology of speech sounds and patterns in the world's natural languages.

Keywords: syllables, onsets, glides, nuclei, codas, Zhangzhou, Southern Min

1 Introduction

Syllable, marked as small Greek sigma σ, has been considered a representational device encompassing principles of segment sequencing (Zec, 2007), although there is no universal agreement upon what a syllable is (Roach, 2000; Laver, 1994; Gimson, 1980). Specifically, syllable is conventionally defined as a complex unit made up of nuclear and marginal element(s) where the nuclear refers to vowel or other syllabic segment, and the marginal element (s) refers to consonant or non-syllabic segment (Laver, 1994). Syllable is also often seen as a section between two successive points of low sonority, and sequences of syllable display a quasiperiodic rise and fall in sonority, and each repeating portion may be viewed as a sonority cycle (Zec, 2007; Clements, 1988; Szigetvári, 2010). Segments within a given syllable are thus arranged in a predictable pattern: the most sonorous occupies the nucleus position, while the less sonorous occurs towards the margin.

As a linguistic universal, syllable can perform several important functions in linguistics (Zec, 2007; Fudge, 1969; Blevins, 1995). For example, it serves as a domain in which certain phonological processes occur, such as, the feature of nasalization may spread within a syllable. It can be referred to as the anchor of suprasegmental phenomena like stress and tone. Such as, in Sinitic languages and many others in Asia, monosyllabic morphemes are largely assigned with a tone to specify lexical meanings. As well as this, human languages vary considerably in how syllables are encoded in their sound systems. For example, English language allows a (C)(C)(C)V(C)(C)(C)(C) structure, like the words strengths /stɹɛŋkθs/ and sixths /sɪksθs/

have four segments occur at the coda position, while in Fijian and Senufo languages, no coda can be attested phonemically (Zec, 2007; Maddieson, 2013).

Given the importance of syllable in the linguistic study of speech sounds, and the diversity of syllable structures that human beings can exploit in their mental grammar and encode in their language practice, a detailed exploration into how syllables are encoded appears to be the most essential step to understand how a given natural language can be used for communication purposes. This study is thus designed to explore the encoding of syllables in Zhangzhou Southern Min, a Sinitic dialect that is spoken in the southern Fujian province of China. Incorporating the modern phonological and phonetic theories and traditional Sinitic linguistics, this study particularly addresses three major issues: (1) how syllable is constructed at the abstract level? and (2) what constitutes the individual syllable components: onsets, glides, nuclei, and codas? and (3) how segments are realised in phonetics. Attention is also given to explaining why certain phonemes are posited when controversial solutions exist. The exploration has sought to go beyond the inconsistency and inadequacy of earlier studies and is expected to serve as a model for thorough investigations of the sound system across Sinitic dialects using modern phonetic/phonological theories and analytical approaches. It contributes important language data to the generalisation of the areal characteristics within the Sino-Tibetan language family, but also to the typology of sound patterns in world's natural languages, while shedding an important light on how human beings encode syllables in their mental grammar and decode syllables in their language practice.

2 Zhangzhou and syllables

2.1 Zhangzhou

Zhangzhou is a prefecture-level city situated in the Southern Fujian province of mainland China. It covers an area of approximately 12,600 square kilometres and hosts a registered population of about 5.10 million (Huang, 2018; 2020; 2022). The colloquial language spoken by native Zhangzhou people is predominantly Southern Min (ISO 639-3 [nan]), which is anglicized as Hokkien for its colloquial pronunciation for Fujian and referred to as Zhangzhou speech in this study. This speech is mutually intelligible with Southern Min varieties of Quanzhou, Xiamen and Taiwan; but is entirely unintelligible with other Chinese dialects (e.g., Mandarin, Hakka, Cantonese, Wu, Xiang, and Gan). Certain regional variations can be observed among its eleven administrative areas that incoporate Longwen, Xiangcheng, Changtai, Longhai, Zhangpu, Yunxiao, Dongshan, Zhao'an, Nanjing, Hua'an and Pinghe county (Huang, 2018; 2020). It is thus important to recognize the regional variation and clarify the research locality before any linguistic study of this dialect.

Given the regional variation, this study restricts the analysis to the synchronic speech of the urban area of Xiangcheng and Longwen districts, which is conventionally considered to be historically-socially-culturally-linguistically-and-geographically representative of Zhangzhou as an independent Southern Min variety.

2.2 Syllable structure

A template of C(G)V(X) can be generalized to characterize the internal syllable structure of synchronic Zhangzhou data, in which onset (C) and nucleus (V) are obligatory while glide (G) and coda (X) are optional (Huang, 2018; 2020). Prevocalic glides possess an independent status (G), while the postvocalic glides are assigned to be one type of syllable coda (X). The components that constitute nuclei are diverse, comprising oral vowels, nasalized vowels, and syllabic nasals. Similarly, the components that occupy the coda position are also diverse, including post-vocalic glides, nasal consonants, and obstruent codas. The whole segmental inventory of Zhangzhou Southern Min is consisting of 15 onsets, 2 prevocalic glides, 13 nuclei, and 8 codas, as summarized in Table 1. Four syllable types, including CV, CGV, CVC and CGVC, can be generalized, as illustrated in Table 2, in which lexical tones are transcribed using Chao (1930)'s notional system, with 5 representing the highest pitch level and 1 the lowest.

Component		Phoneme
C	onset	p, p^{h}, ɓ, t, t^{h}, ɗ, k, k^{h}, ɠ, ts, ts^{h}, s, z, ħ, ʔ
G	glide	j, w
V	nucleus	i, e, ɛ, ɐ, ɔ, ɵ, u, ĩ, ɛ̃, ɐ̃, ɔ̃, m, ŋ
X	coda	j, w, m, n, ŋ, p, t, k

Table 1: Summary of Zhangzhou segments.

Type	Example 1	Example 2
CV	tĩ35 ‘sweet’	ku33 ‘worn’
CGV	sjɐ51 ‘write’	kwɐ35 ‘song’
CVC	sim35 ‘heart’	kɐw41 ‘arrive’
CGVC	tsjɐp41 ‘juice’	kjɐm22 ‘salty’

Table 2: Examples of Zhangzhou syllable types

2.3 Linguistic data

The empirical data upon which this study is built are collected by the author from multiple field trips from 21 native speakers (9 males and 21 females) from the urban districts of Longwen and Xiangcheng. They are selected based on a set of criteria including intellectual curiosity, physical condition, linguistic environment at home, age, education background, and competence of other dialects and/or languages, with an average age of 56.5 for males, and 50 for females. The collected data are acoustically processed to present their acoustic property with respect to a particular category. Specifically, the first three formant patterns are extracted to demonstrate the contrastive distribution of vowels, and to examine the existence of obstruent codas in this dialect (Huang, 2023). The values of voice onset time (VOT) (Abramson & Whalen, 2017) are extracted to examine the distribution of plosive sounds. Normalization is conducted to reduce speaker-dependent variation. For instance, the Lobanov (1971)’s approach is used for vowel quality and obstruent coda examination, as shown in (1).

$$F_iN = (F_i - M_i)/\sigma_i \quad (1)$$

3 Zhangzhou Onsets

Onset is the consonant sound (s) that occur (s) at the beginning of a syllable, typically before the nucleus (Zec, 2007; Fudge, 1969; Blevins, 1995). Zhangzhou possesses 15 onset phonemes (/p, pʰ, ɓ, t, tʰ, ɗ, k, kʰ, ɠ, ts, tsʰ, s, z, ħ, ʔ/) that are contrastive to distinguish lexical items, as illustrated in Table 3 with (near-) minimal pairs. The onsets can be characterized in several different ways. As seen, they can be classified into five categories in terms of the place of articulation, including bilabial /p, pʰ, ɓ/, alveolar /t, tʰ, ɗ, s, z, ts, tsʰ/, velar /k, kʰ, ɠ/, pharyngeal /ħ/, and glottal /ʔ/, and into three categories in terms of the manner of articulation, including plosive (/p, pʰ, ɓ, t, tʰ, ɗ, k, kʰ, ɠ, ʔ/), fricative (/s, z, ħ/), and affricate (/ts, tsʰ/). The alveolar sounds occupy more than 45% of the total onsets, while the occlusive sounds occupy up to 67% of the onset inventory. Only two phonemes (/ts, tsʰ/) are affricates that are exclusively alveolars and contrasted in aspiration. Three fricatives (/s, z, ħ/) can be identified over the alveolar and glottal positions, with the two alveolars further differing in voicing.

The voicing and aspiration are both have a phonemic function in this dialect, which can classify the nine supraglottal plosives into voiceless unaspirated (/p, t, k/), voiceless aspirated (/pʰ, tʰ, kʰ/), and voiced categories (/ɓ, ɗ, ɠ/). The special airstream mechanism of ingressive glottalic, known as implosive, is employed phonemically (Huang, 2021; Huang & Hyslop, 2022). Three implosives (/ɓ, ɗ, ɠ/), differing in the place of articulation, can be identified for those tokens that were previously transcribed as ordinary egressive pulmonic consonants (/b, l, g/) (Dong 1959; Lin 1992; Ma 1994; FJG 1998; ZZG 1999; Zhou 2006).

C	Example 1	Example 2
/p/	pi51 ‘compare’	piŋ22 ‘friend’
/pʰ/	pʰi51‘scab’	pʰiŋ22‘comment’
/ɓ/	ɓi51 ‘rice’	ɓiŋ22 ‘bright’
/t/	ti51 ‘resist’	tiŋ22 ‘pavilion’
/tʰ/	tʰi51 ‘store’	tʰiŋ22 ‘stop; suspend’
/ɗ/	ɗi51 ‘you’	ɗiŋ22 ‘zero’
/k/	ki51 ‘point out’	kiŋ22 ‘lift up; engine’
/kʰ/	kʰi51 ‘tooth’	kʰiŋ22 ‘jade’
/ɠ/	ɠi51 ‘language’	ɠiŋ22 ‘welcome’
/ts/	tsi51 ‘cook’	tsiŋ22 ‘emotion’
/tsʰ/	tsʰi51 ‘mouse’	tsʰiŋ22 ‘banyan tree’
/s/	si51 ‘die’	siŋ22 ‘complete’
/z/	zi51 ‘a quantifier’	zin 22 ‘people’
/ħ/	ħi51 ‘happy’	ħiŋ22 ‘shape’
/ʔ/	ʔi51 ‘chair’	ʔiŋ22 ‘glory; honor’

Table 3: Examples of Zhangzhou onsets

The three-way contrast of supraglottal plosives in Zhangzhou can be well distinguished in terms of the voice onset time (VOT), which is defined as the time between the release of the oral constriction for plosive production and the onset of vocal fold vibration (e.g., Abramson & Whalen, 2017). Figure 1 plots the VOT distribution of plosive sounds which is derived from quantifying 1147 samples (=6 tokens * 9 plosives * 21 speakers) based on the empirical data that the first author collected in the urban districts of Zhangzhou city in 2015. As seen, the voiceless unaspirated plosives (/p, t, k/), regardless of the place of articulation, consistently show positive values slightly above zero from 0.013 ms to 0.024 ms, because the vocal folds vibrate immediately after the oral constriction is released. The voiceless aspirated stops (/pʰ, tʰ, kʰ/) show steep positive values from 0.067 ms to 0.083 ms because, after the plosive constriction releases, there is a period for the articulation of aspiration, causing a delay in the onset of vocal fold vibration. Contrastively, the implosives (/ɓ, ɗ, ʛ/) present steep negative values between -0.077 ms and -0.064 ms because the vocal folds vibrate before the oral constriction is released (Huang & Hslop 2022).

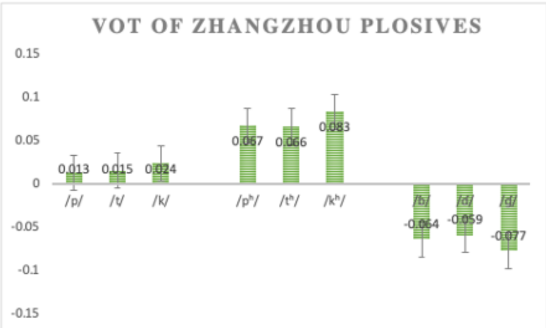


Figure 1: VOT of Zhangzhou onsets.

4 Zhangzhou Glides

Two prevocalic glides /j/ and /w/ are contrastively used in the Zhangzhou speech, as illustrated in Table 4. Phototactically, their occurrence is subject to the OCP (obligatory contour principle) on the place of articulation. Rhymes like /*wɵ/ and /*wɔ/ are prohibited because of their sharing common feature in [+labial], whilst rhymes like /jɵ/ and /wɐ/ are attested. Phonetically, the gliding features [+palatal] and [+labial] are continuous, which can substantially affect the phonetics of preceding onsets. For example, the implosives /ɓ/ and /ʛ/ are realised as labialised and voiced fricatives [βʷ] and [ɣʷ], respectively, before the labial glide /w/, while the alveolar occlusives /t/, /tʰ/, and /ɗ/ are dentalised and become [t̪], [t̪ʰ], and [ɗ̪] before the palatal glide /j/ (Huang & Hyslop, 2022; Huang, 2023).

Glide	Example 1	Example 2
/j/	kjɐ41 'post'	kjẽ 35 'capital city'
/w/	kwɐ41 'hang'	kwẽ35 'liver'

Table 4: Examples of Zhangzhou glides.

5 Zhangzhou Nuclei

Individual syllables are viewed containing crest and trough (or peak and valley) of sonority because of the opening and closing action of the jaw during the production (Zec, 2007; Clements, 1988). The segment occupying the crest/peak of the sonority is called nucleus which is compulsory to occur in speech sounds. In Zhangzhou Southern Min, the segments that can fill in the nucleus position are diverse, comprising seven oral vowels, four nasalized vowels, and two nasal consonants.

5.1 Oral nuclei

Seven oral vowel phonemes (/i, u, e, ɛ, ɐ, ɵ, ɔ/) can be identified in the empirical speech, as illustrated in Table 5 with minimal pairs. They present a three-way contrast of High (/i, u/), Mid (/e, ɵ/), and Low (/ɛ, ɐ, ɔ/) in terms of tongue height, and also a three-way contrast of Front (/i, e, ɛ/), Central (/ɐ, ɵ/), and Back (/u, ɔ/) in terms of the backness of tongue in the oral cavity. Likewise, a two-way contrast of rounded (/u, ɵ, ɔ/) and unrounded (/i, e, ɛ, ɐ/) can be figured out as to the rounding status of tongue lips over the articulation.

V	Example 1	Example 2
/i/	ki35 'base'	si35 'poetry'
/e/	ke35 'chicken'	se35 'comb'
/ɛ/	kɛ35 'family'	sɛ35 'yarn; gauze'
/ɐ/	kɐ35 'glue'	sɐ35 'grasp'
/ɔ/	kɔ35 'mushroom'	sɔ35 'crisp; biscuit'
/ɵ/	kɵ35 'cake'	sɵ35 'rub; knead'
/u/	ku35 'tortoise'	su35 'lose;

Table 5: Examples of Zhangzhou oral vowels.

The phonemic contrast of Zhangzhou oral vowels can also be differentiated in their acoustic distribution. Figure 2 plots their acoustic distribution which is derived by Lobanov-normalising (1971) 15876 (=6 tokens* 7vowels*9 sampling points*2 formants*21 speakers) values based on the field data that Huang (2018) collected from 21 native speakers. As seen, the front high vowel [i] presents the largest the second formant value (F2) but its first formant value (F1) is the smallest, while the back high vowel [u] has a similar F2 as the vowel [i], but its F1 is the largest. In other words, the more front the vowel, the higher the second formant (F2), while the lower the vowel, the larger the first formant (F1).

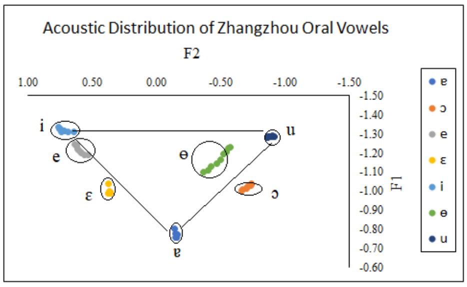


Figure 2: Acoustics of Zhangzhou oral vowels.

5.2 Nasalised nuclei

The feature of nasality plays a contrastive function in the nucleus system of Zhangzhou speech. Four out of the seven oral vowels can be nasalised (/ĩ, ɛ̃, ɐ̃, ɔ̃/) and used phonemically (e.g., */ti35/ 'pig'*/ vs. */tĩ35/ 'sweet'*), as illustrated in Table 6. This suggests a contrastive function between a lowered velum and a raised velum in speech production of this dialect. The number of contrastive nasal vowels (four) is apparently much fewer than that of oral vowels (seven), but it reflects a cross-linguistic tendency about the distribution of nasalised vowels in human languages. For example, Yoruba has seven contrastive oral vowels, but only five nasalised vowels are documented; Maba has twelve oral vowels, but only one nasalised vowel /ũ/ is reported (Hajek 2005).

Nuclei	Example 1	Example 2
/i/-/ĩ/	ti35 'pig'	ki35 'branch'
	tĩ35 'sweet'	kĩ35 'alkali'
/ɛ/-/ɛ̃/	pɛ33 'father'	kɛ35 'home'
	pɛ̃33 'disease'	kɛ̃35 'net'
/ɐ/-/ɐ̃/	p^{h}ɐ41 'beat'	kɐ35 'glue'
	p^{h}ɐ̃41 'puff'	kɐ̃35 'prison'
/ɔ/-/ɔ̃/	kɔ22 'paste'	ɦɔ51 'tiger'
	kɔ̃22 'snore'	ɦɔ̃51 'well'

Table 6: Examples of Zhangzhou nasalised nuclei.

The phonemic contrast between oral and nasalised vowels in Zhangzhou can also be justified in acoustics. Figure 3 shows the acoustic distribution of the four oral-nasalised vowel pairs (Huang 2022). As indicated, the nasalised vowels tend to have larger values in both F1 and F2 than their oral counterparts, except the nasalised vowel [ɐ̃] has a smaller F2 value than the oral vowel [ɐ]. Such a manifestation suggests that, during the articulation of related tokens, the tongues of native speakers move to a more fronting but lower position, reflecting the acoustic coupling between oral cavity and nasal cavity that is evoked by the lowering of the velum for the airstream to pass through both cavities (Hawkins & Stevens 1985).

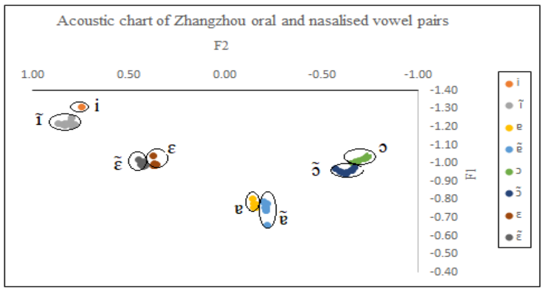


Figure 3: Acoustics of oral vs. nasalized vowels.

5.3 Syllabic nasals

Apart from the oral and nasalised vowels, two nasal consonants /m/ and /ŋ/ can also occupy the nucleus position in syllables and perform a distinctive function in semantics in this dialect. For example, the bilabial syllabic nasal /m̩/ can carry five different pitch contours to differentiate five lexical meanings, such as /ʔm24/ 'drink', /ʔm22/ 'flower bud', /ʔm51/ 'aunt', /ʔm41/ 'affirmative' and /ʔm33/ 'negative'. The velar nasal /ŋ̩/ can follow onset consonants of different places and manners of articulation to construct lexical morphemes, e.g., /t^{h}ŋ35/ 'soup', /tŋ22/ 'long', /sŋ51/ 'play', /kŋ41/ 'steel', /k^{h}ŋ35/ 'bran' and /tsŋ33/ 'statement'. This supports the cross-linguistic syllabification algorithm that syllable nucleus is not absolutely a projection of vowel; instead, the nuclei can be consonantal (Huang 2022).

However, the two nasals turn out to occur in the place where present accidental gaps for nasalised vowels /*ũ/ and /*ɵ̃/, as reflected in Table 7. It thus seems logically plausible to posit their underlying forms as /ũ/ and /ɵ̃/. The reason to propose /m̩/ and /ŋ̩/ as contrastive phonemes is mainly because of their high productivity in colloquial conversations. For example, according to the calculating result on Huang's rhyme tables (Huang 2019), the velar nasal /ŋ/ can combine with 12 out of 15 onset phonemes to create monosyllabic morphemes, while being able to occur in 6 out of 8 lexical tones. In contrast, no tokens have the nasalised vowels [ũ] and [ɵ̃]. Another important consideration is to make this Southern Min variety comparable with other Sinitic dialects, because the phonemic status of syllabic nasals is conventionally acknowledged as a typical feature of Southern Chinese dialects (Huang 2022).

Oral	Nasalised	Syllabic nasal
i	ĩ	*
u	*	m̩
e	*	*
ɵ	*	ŋ̩
ɛ	ɛ̃	*
ɐ	ɐ̃	*
ɔ	ɔ̃	*

Table 7: Asymmetrical distribution of Zhangzhou nucleus types.

6 Zhangzhou Codas

The coda, known as auslaut, comprises a principal subpart of a syllable that follows a nucleus, and forms a larger constituent of rhyme with the nucleus (Zec 2007). In Zhangzhou Southern Min, the codas can be consonantal that include nasals and obstruent stopes, but also can be glides. They are optional to occur but are prohibited from forming consonantal clusters in syllables.

6.1 Glide codas

Two postvocalic glides /j/ and /w/ can contrast lexical meanings in Zhangzhou, such as the minimal pair of /kɐj41/ 'boundary' and /kɐw41/ 'enough', as illustrated in Table 8. However, phototactically, their occurrences are severely constrained. They can only occur after /ɐ/ and /ɐ̃/, so that the eight glide-ending rhymes of /ɐw/, /ɐ̃w/, /ɐj/, /ɐ̃j/, /jɐw/, /jɐ̃w/, /wɐj/, and /wɐ̃j/ are attestable. In the meanwhile, they are prohibited to occur after the other nine nuclei, thus, rhymes like /*ɛj/, /*ɛw/, /*ej/, /*ew/, /*ɔj/, /*ɔw/, /*ɛ̃j/, and /*ɔ̃j/ are not attested in the data.

Coda	Example 1	Example 2
/j/	kɐj41 'boundary'	ʔɐj41 'love; like'
/w/	kɐw41 'enough'	ʔɐw41 'vexed'

Table 8: Examples of Zhangzhou glide codas.

Phonologically, they are considered as one type of syllable coda (C) rather than as a part of the nucleus (V) and as an independent syllable component (G), because of two considerations. First, as mentioned above, they are less productive to create syllable rhymes and lexical items. Additionally, this treatment can simplify the sub-syllabic modelling for the tokens that only occupy a small number of the vocabulary. Because, given them an independent phonological status G, there logically may have a maximal syllable template C(G)V(G)(C), which contains as many as five constituents. In the contrast, assuming them as one type of syllable codas, the syllable template can be simplified as C(G)V(C), with the types of CVG and CGVG eliminated, enabling the phonological analysis to be simple, economical, and elegant. Therefore, positing them as one type of syllable coda (X) not only respects the reality of their low productivity but also can simplify the phonological analysis.

6.2 Nasal codas

Zhangzhou has three nasal codas /m/, /n/, and /ŋ/ that differ in the place of articulation to distinguish lexical items, as illustrated in Table 9. They are found to be productive in formulating syllables. On the basis of the calculating result on the homonym inventory (Huang, 2019), 421 syllables have the CVN type, while 216 have the CGVN type, the total of which occupy near 30% of the total of the syllables that are colloquially attested.

Coda	Example 1	Example 2
/m/	kɐm41 'supervise'	ʔɐm41 'dim'
/n/	kɐn41 'separate'	ʔɐn41 'file, proposal'
/ŋ/	kɐŋ41 'descend'	ʔɐŋ41 'earthen jar'

Table 9: Examples of Zhangzhou nasal codas.

The three-way nasal contrast in the coda position demonstrates Zhangzhou as a typical dialect spoken in Southern China. Because many dialects in the North are often reported either losing some nasal codas or having them merged over their diachronic development. For example, the bilabial coda /m/ has been lost in Mandarin, Xiang and Gan dialects, while other dialects like Jin, Wu, and Eastern Min only preserve the vela coda /ŋ/ (Chen & Anderson, 2011).

6.3 Obstruent codas

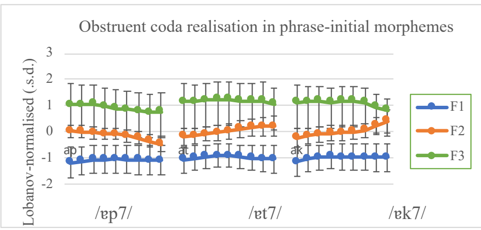
Phonemically, this dialect possesses three obstruent codas (/p, t, k/) that differ in the place of articulation to distinguish lexical items, as illustrated in Table 10. The codas pattern with their nasal counterparts to classify Zhangzhou syllables and tones into stopped and unstopped classes. The stopped category refers to those syllables ending in obstruent sounds, while the unstopped category refers to those ending in sonorant segments.

Table 10: Examples of Zhangzhou obstruent codas.

X	Example 1	Example 2
/p/	kɐp41 'pigeon'	ʔɐp41 'restrain;
/t/	kɐt41 'tie; knot'	ʔɐt41 'suppress'
/k/	kɐk41 'horn'	ʔɐk41 'irrigate;

Phonetically, the realisation of obstruent codas appears to be contextually sensitive: they are not realised in the utterance-final context but are realised in the non-final context. The changing formants for different obstruent coda articulation can also be justified by quantifying 21 speaker's utterances, as plotted in Figures 4 and 5, each of which is derived from Lobanov-normalising 5103 (=3 tokens * 9 sampling points * 3 formants * 3 obstruent codas * 21 speakers) formant values (Huang, 2023).

Figure 4: The non-realisation of obstruent codas in the utterance-final context.



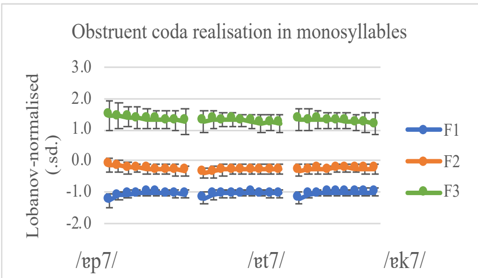


Figure 5: The realization of obstruent codas in the non-utterance-final context.

In Figure 4, the monosyllabic samples share the same tone 7 and same vowel /ɐ/ at the underlying level but differ in the type of obstruent codas. As indicated, their normalized formant patterns are consistently stable without significant change over the major rhyme portions (/ɐp7/, /ɐt7/, /ɐk7/), regardless of the place of articulation for their underlying obstruent codas. The manifestation suggests that (a) the related vowel is a monophthong, and (b) no obstruent coda is articulated. In other words, during the articulation, the related active articulator is not manipulated to form a complete oral constriction with the passive articulator at a particular point along the vocal tract (Huang, 2023). This justifies the non-realisation of obstruent codas in the utterance-final context.

Figure 5 models the normalized formant patterns of the first morphemes of disyllabic samples, which also share the same vowel /ɐ/ and the same tone 7 but differ in the type of obstruent codas. As seen, the formant patterns are different across the three rhymes, with the F2 curve transiting to three different targets. In the phrase-initial rhyme /ɐp7/, the F2 curve shows a falling tendency to 0.5 s.d. below the mid, indicating a postvocalic-bilabial obstruent coda. In the phrase-initial rhyme /ɐt7/, the F2 curve is raised up from -0.2 towards a point that is 0.2 s.d. above the mid, indicating an alveolar place of articulation for subsequent obstruent sound. In the phrase-initial rhyme /ɐk7/, the F2 curve noticeably increases while the F3 curve abruptly drops, but the two curves tend to converge at a point around 0.5 s.d., showing a characteristic of velar pinch that indicates the production of a velar obstruent coda. Thus, the changing formant transitions toward different locus frequencies sufficiently justify the existence of obstruent codas of different places of articulation in this non-utterance-final context (Huang, 2023).

6.4 Glottal stop

A contrastive glottal stop coda /ʔ/ is documented in prior studies of Zhangzhou (Lin 1992; Ma 1994; ZZG 1999; Dong 1959; FJG 1998; Gao 1999), which can follow different nuclei to form stopped syllables. For example, Ma (1994) documents 24 glottal stop-ending rhymes that include /iʔ/, /uʔ/, /ɛʔ/, /oʔ/, /eʔ/, /m̩ʔ/, /aʔ/, /ɔʔ/, /ĩʔ/, /ɔ̃ʔ/, /ãʔ/, /ɛ̃ʔ/, /ŋ̍ʔ/, /iaʔ/, /iãʔ/, /ioʔ/, /iɔʔ/, /uaʔ/, /uãʔ/, /ueʔ/, /auʔ/, /ãuʔ/, /iauʔ/, /iãuʔ/, and /uãiʔ/. However, contradicting prior convention, this study, in accordance with Huang's (2018; 2020) research, declines a phonemic status of glottal stop coda. The main reason for this treatment is that the glottal stop tends to disappear synchronically, and the related syllables become open without a stop coda that can be perceived overtly. This disappearing of glottal stop coda causes related syllables to have different phonetic manifestations from those syllables that are conventionally transcribed having the same tones but ending in supra-glottal obstruent codas. Specifically, for those obstruent-ending syllables, they have a shorter duration and a depressed tonal pitch, because of the laryngealisation that occur on rhyme portions (Huang, 2023), which is also cross-linguistically reported in tonal languages (Pan, 2017; Ladefoged & Maddieson, 1996; Garellek, 2013; Kuang, 2011). For those syllables whose glottal stop has been deleted, they turn out to have a longer duration and the related tonal pitch contour is shown to be stable without any depressing effect.

Additionally, as a direct consequence of the synchronic disappearance of the glottal stop, a new tone (tone 8) emerges that gives rise to a system of eight tones in the present-day Zhangzhou speech, contradictory to a seven-tonal system in previous documentations (Huang, 2018; 2020). Tone 8 is associated with the historically assumed glottal stop coda while tone 7 is related to obstruent codas. In whichever the linguistic context: citation or sandhi position, tone 8 presents different manifestations from tone 7. Specifically, tone 8 presents a low level [22] with an extra-long duration in citation, whereas tone 7 presents a low-level with a falling trend [221]. Likewise, in the non-rightmost position, which is also referred to as sandhi position, tone 8 presents a low-falling [32] contour with a medium duration, whereas tone 7 that are associated with obstruent codas shows a low-falling [32] contour with an extra-short duration. This can be seen in Figure 6 that plots the normalized F0 contours of tone 7 (left) and tone 8 (right) in both citation (up) and sandhi (bottom) contexts (Huang, 2018; 2020).

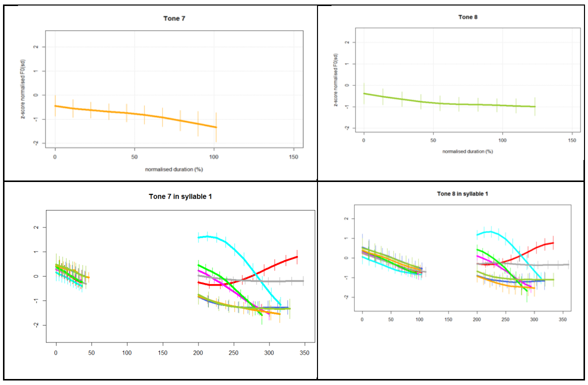


Figure 6: The F0 contours of tone7 and 7 in citation and sandhi contexts.

On the contrary, if the glottal stop coda is posited as being phonemic to maintain the historical convention, the treatment not only could create substantial redundancy in the construction of rhyme inventory and dictionary but could result in considerable number of unnecessary complexities for phonotactic analyses and interpretations. For example, more than 28% of the overall rhyme inventory are documented ending in the glottal stop in previous studies (Ma, 1994; FJG, 1998; ZZG, 1999). Thus, the proposal of declining a phonemic status of glottal stop coda not only reflects and respects the phonetic reality but also maintaining the phonological analysis to be elegant, simple, and economic.

7 Conclusion

This study explores how syllables are encoded in Zhangzhou Southern Min, a Sinitic dialect spoken in South Fujian of mainland China. This dialect possesses a C(G)V(X) syllable structure, in which onset and nucleus are compulsory while the prevocalic glide and coda are optional to occur. Prevocalic glides possess an independent status, while the postvocalic glides are assigned to be one type of syllable coda. Nuclei comprise oral vowels, nasalized vowels, and syllabic nasals, while the codas incorporate glides, nasal consonants, and obstruent codas. A segmental system of 15 onsets, 2 prevocalic glides, 13 nuclei, and 6 codas can be posited, whose components are contrastively used to distinguish lexical items.

The study establishes a comprehensive picture of the encoding of syllables in Zhangzhou Southern Min, which substantially advances and stretches our understanding of its segmental phonetic and phonology. It is of both theoretical and empirical significance to serve as a model for a thorough investigation of the sound system in Sinitic dialects from the perspectives of modern linguistic theory and analytical approaches. The results go beyond the inconsistence and inadequacy of prior studies that are auditory-based and lack of scientific support. The study sheds an important light on the generalization of the areal characteristic of syllables within the Sino-Tibetan language family., while contributing vital empirical data to the typology of sound patterns in world's natural languages. It also enlightens our knowledge of how human beings encode syllables in their mental grammar and decode syllables in language practices across different speech communities.

8 Limitations

This study mainly discusses the categorical distinction of segments that can occur at different positions within syllables. It presents an overall profile of the segmental system of this dialect.

However, studies need to be conducted as to how individual phonemes are realised at the phonetic level and whether their realisations may be affected by their subsequent sounds in connected speech. Because the speech organs of human beings are supposed to operate in a continuous motion, some phonemes are expected to have allophonic variants that may be in complementary distribution in real-world utterances. Likewise, human beings may initiate different laryngeal settings over the articulation, thus, some sounds may be produced with different voice quality or different airstream mechanism. These are of linguistic significance to understand the nature of speech sounds in a given natural language, but they are not presented in this study because of the layout issue and the constriction of research scope.

References

Alfred Gimson. 1980. *An Introduction to the Pronunciation of English (3rd)*. London, United Kingdom: Edward Arnold.

Arthur S Abramson and Douglas H Whalen. 2017. Voice onset time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, 63, 75–86.

Baoqin Lin. 1992. Zhangzhou fangyan cihui (Zhangzhou vocabularies). *Fangyan*, 1-3.

Boris M Lobanov. 1971. Classification of Russian Vowels Spoken by Different Speakers. *The Journal of the Acoustical Society of America*, 49(2B), 606-608.

Chongqi Ma. 1994. *Zhangzhou fangyan yanjiu (Studies of Zhangzhou dialect)*. Hongkong: Zongheng Chubanshe.

Draga Zec. 2007. The syllable. In P.D. Lacy. *The Cambridge Handbook of Phonology*,161-194. United Kingdom: Cambridge University Press.

Eric C Fudge. 1969. Syllables. *Journal of Linguistics*, 5(2), 253-286.

FJG. 1998. *Fujianshengzhi fangyanzhi (Fujian Province Gazette-Dialect Volume)*. Beijing: Fangzhi Chubanshe.

George Clements. 1988. The Role of the Sonority Cycle in Core Syllabification. *Working Papers of the Cornell Phonetics Laboratory*, No. 2. Ithaca, New York: Cornell University.

Ho-hsien Pan. 2017. Glottalization of Taiwan Min checked tones. *Journal of the International Phonetic Association*, 47(1), 37-63.

Ian Maddieson. 2013. Syllable Structure. In: D. Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Jianjing Kuang. 2011. Production and Perception Maps of the Multidimensional Register Contrast in Yi. *UCLA Working Papers: Phonetics*, 1-30.

John Hajek. 2005.Vowel Nasalization. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary.

John Laver. 1994. *Principles of Phonetics. Cambridge*, United Kingdom: Cambridge University Press.

Juliette Blevins. 1995. The Syllable in Phonological Theory. In J. Goldsmith, *Handbook of Phonological Theory*. 206-255. London, United Kingdom: Basil Blackwell, London.

Marc Garellek. 2013. *Production and Perception of Glottal Stops*. Doctoral dissertation, UCLA. University of California Los Angeles. Retrieved from https://escholarship.org/uc/item/7zk830cm.

Peter Ladefoged and Ian Maddieson. 1996. *The Sounds of the World' s Languages.* Oxford: Blackwell.

Peter Roach. 2000. *English Phonetics and Phonology.* Cambridge, United Kingdom: Cambridge University Press.

Peter. Szigetvári. 2010. Syllables. In N. Kula. E.D., B. Botma., & K. Nasukawa (eds). *Bloomsbury Companion to Phonology*, 64-94. London, United Kingdom: Bloomsbury Companions.

Ran Gao. 1999. Introduction to the sound system of Zhangzho. In *Minnan Dialect-Studies of Zhangzhou Variety*.109-116. Beijing, China: Zhongguo Wenlian Chubanshe.

Sarah Hawkins and Kenneth. N Stevens. 1985. Acoustic and perceptual correlates of the nonnasal-nasal distinction for vowels, *Journal of the Acoustical Society of America*, (77), 1560–1575.

Tonghe Dong.1959. *Si ge Minnan fangyan (Four Southern Min varieties)*. Taipei: Zhongyang Yanjiuyuan.

Ying Chen and Susan Guion Anderson. 2011. Perceptual confusability of word-final nasals in Southern Min and Mandarin: Implications for coda nasal mergers in Chinese. *International Congress of Phonetic Sciences*. Hongkong, August 17-21.

Yishan Huang and Gwendolyn Hyslop 2022; submitted. Zhangzhou Implosives and Their Variations. In *Proceedings of the 20th Annual Workshop of the Australasian Language Technology Association (ALTA 2022)*, 14-16 December 2022, Flinders University, Adelaide, Australia

Yishan Huang. 2018. *Tones in Zhangzhou: Pitch and Beyond.* Doctoral dissertation: The Australian National University.

Yishan Huang. 2019. *Zhangzhou Southern Min: Rhyme tables, Homonyms, Heteronyms, and Vernacular documentation*. Munich, Lincom Europa

Yishan Huang. 2022. Nasality in Zhangzhou: Distribution and Constraint. In *Proceedings of the 25th Conference of the Oriental COCOSDA (Coordination and Standardisation of Speech Databases and Assessment Techniques)*, 24-26 November 2022; Hanoi University of Science – Vietnam National University of Hanoi, Hanoi, Vietnam

Yishan Huang. 2022. Right-dominant tones in Zhangzhou: on and through the phonetic surface. In *Proceeding of the 34th Annual Conference on Computational Linguistics and Speech Processing in Taiwan (ROCLING)*, 21-22 November 2022, Taipei Medical University, Taiwan, China.

Yishan Huang. 2023. (accepted). Obstruent Codas in Zhangzhou Southern Min: Contexual Realisations and Interrelated Effects. *Journal of Chinese Phonetics*.

Yuanren Chao. 1930. ə sɪstəm əv “toun-lɛtəz” (A system of “tone letters”). *Le Maître Phonétique*, 45, 24-27.

ZZG. 1999. Zhangzhou Shizhi Fangyan (Zhangzhou chorography-dialect). Vol. 49. Beijing: Zhongguo Shehui Kexue Chubanshe.