Encoding Tone Sandhi in Zhangzhou Southern Min
An Inter-disciplinary Exploration

Yishan Huang
1The University of Sydney
yishan.huang@sydney.edu.au

Abstract
This study adopts an interdisciplinary approach to explore the nature of tone sandhi in Zhangzhou Southern Min, a Sinitic dialect spoken in southern Fujian, mainland China. This dialect presents a typical right-dominant tone sandhi system, but the sandhi domain is not phonologically specified, but rather relevant to the boundary of syntactic phrases. The sandhi process is phonologically inert to the category of the surrounding tones; however, it is highly sensitive to the phonetics of their environments. Tonal neutralization occurs across linguistic contexts, making it difficult to determine the directionality of tonal alternation. Thus, this study claims a morphological nature to characterize Zhangzhou tones. The relationship between its sandhi (non-phrase-final) and citation tonal forms is morphophonemic, whereas it is allophonic between phrase-final and citation tones. This exploration substantially stretches and advances our knowledge of tone sandhi as an important language phenomenon in Southern Chinese dialects. It is hoped that it will serve as a model to investigate the nature of tone sandhi in relevant languages.

Index Terms: tone sandhi, interdisciplinary, Zhangzhou

1. Introduction
The realization of tones can be alternated when they come into contact with one another in connected speech. The process of contextually triggered tonal alternation is referred to as tone sandhi in the literature [1-2]. Among those languages where tone sandhi is prevalent, they vary considerably in the way that tones change, what has motivated tones to change, and under what domain tones are supposed to change. For example, Sinitic languages have been conventionally classified as either right-dominant or left-dominant depending on where syllables retain their corresponding tonal forms in citation [1]. The encoding of tone sandhi in real-world utterances may be more complicated than expected. For example, the Southern Min (SM) has long been asserted to exhibit an intricate tone sandhi phenomenon. One tone’s citation form is asserted to be the sandhi form of the other tone. The whole sandhi process among those tones associated with sonorant-ending syllables is considered to form a circular change, as shown in Figure 1 for the Xiamen SM (left) [3] and Taiwan SM (right) [4].

Figure 1. SM tone circle in Xiamen (left) and Taiwan (right).

A number of proposals have been made to explain why tones change in a circular manner and what mechanisms have motivated the circular change in Southern Min. They are created from diverse perspectives, including generative phonology [5], autosegmental phonology [6-7], optimality theory [8-10], psycholinguistic theory [11-12], and lexicalized phrasal phonology [13-15]. However, no agreement was reached regarding a unified account. It is still unclear what tone sandhi is in Southern Min, and knowledge still needs to be improved in terms of the most basic aspects relating to the encoding of tone sandhi.

Driven by this intriguing speech phenomenon of Southern Min tone sandhi, this study, built on new empirical data from Zhangzhou Southern Min, adopts an interdisciplinary approach to address six specific questions: (a) What is the dominance of tone sandhi in Zhangzhou (ZZ)? (b) In what domain is the Sandhi process supposed to operate? (c) How are the tones realized in different positions within the sandhi domain? (d) Whether tonal realizations are affected by surrounding tones. If so, what may condition the variation? (e) What are the relations between sandhi tones and their corresponding citation tones, and also between non-sandhi tones and citation tones? and (f) How do native speakers model various tonal forms in their mental grammar? The scientifically grounded exploration is supposed to stretch and advance our knowledge of tones and serve as a model to investigate tone sandhi as an important language phenomenon in Southern Chinese dialects and other tonal languages where relevant.

2. Research Design

2.1. Zhangzhou and Speech
Zhangzhou is a prefecture-level city situated in the southern Fujian province of Southern China, covering an area of approximately 12,600 square kilometers and a population of approximately 5.10 million population [16-17]. The colloquial language spoken by native speakers is Southern Min, also known as Hokkien, which is mutually intelligible with Southern Min varieties of Xiamen, Quanzhou, and Taiwan, but is unintelligible to other Chinese dialects, such as Mandarin, Hakka, Wu, and Cantonese. Certain regional variations can be observed in the sound system of the 11 administrative areas [18-19]. This study thus restricts the locality to the urban areas of Longwen and Xiangcheng districts, which are considered to be historically, socially, socially, culturally, and linguistically representative of Zhangzhou [18].

2.2. Data Elicitation
The data used in this study were obtained from a formal linguistic field trip in 2015 [16-17]. Twenty-one native
speakers, including 9 males and 12 females from the urban areas of Longwen and Xiangcheng districts, participated in the field data recording in an acoustically absorbent room at Zhangzhou Hotel. The average age of the males was 56.5 and 50 for females.

The corpus used in this study incorporated approximately 160 monosyllabic morphemes and 588 disyllabic phrases. Tokens were chosen across different syllable types: onsets of different manners and places of articulation, and vowels of varying height and frontness. This design maximally balanced the intrinsic perturbation effects on tonal F0 from the tautosyllabic segments. The tokens to be elicited were all written in simplified Chinese characters and presented via PowerPoint, with one token on one slide, which substantially ensured a balanced intensity and well-controlled speech rate over the recording [16-17]. All recordings were digitized at a sampling frequency of 44100 Hz in Praat.

2.3. Acoustic Processing

Acoustic processing of the obtained field data was conducted using Praat. Tonally relevant durations were identified by incorporating all elements except syllable onsets [16-17]. Praat scripts were run to automatically extract tonal durations and F0 values. F0 was extracted from ten equipartition sampling points. Because the acoustic signals were highly variable that carried both linguistic and extralinguistic information [20-21], all extracted acoustic data were normalized to reduce both inter- and intra-speaker variation. Formula (1) is the z-score approach for F0, and (2) is the absolute approach for duration [16-17], [20-21].

\[ Z_i = (X_i - m)/s \]  \hspace{1cm} (1)
\[ D_{norm} = (D - D_{mean})/100 \]  \hspace{1cm} (2)

In these formulas, the parameters \( m \) and \( s \) represent the raw mean value and the standard deviation estimated from all sampling values over all tokens from a given speaker, respectively. \( X_i \) is an observed value, while \( Z_i \) refers to its normalized value. \( D \) is the observed duration value and \( D_{norm} \) is its corresponding normalized value, expressed as a percentage of the average duration of all tones from the speaker being considered.

2.4. Statistical Testing

This study applied the pairwise t-test by effect size to examine whether the tonal realizations in disyllabic combinations are affected by their surrounding tones in a statistically significant way [16-17], [22]. If yes, to what extent are they affected, and what conditions do these variations occur? This testing required all possible pairwise comparisons of the values derived from quantifying and normalizing the acoustic utterances of the 21 speakers. For example, as illustrated in Figure 2, 28 (8 × 7/2) paired normalized F0 differences were tested to examine whether F0 onset of tone 1 was affected by its preceding tones.

3. Zhangzhou Citation Tones

Zhangzhou has received extensive documentation of its monosyllabic citation tones [18-19], [23-27]. However, prior studies before Huang [16-17]’s initiatives are dominantly impressionistic and identify a seven-way tonal contrast with inconsistent transcriptions. For example, tone 2, which corresponds to the Yangping tone in terms of the Middle Chinese tonal category, was transcribed in five different ways: [212], [13], [12], [23], and [22], which cover dipping, rising, and level contour shapes [28]. Huang [16-17] advocates an eight-tone system based on two important facts that (a) Zhangzhou tones are multidimensional, and (b) Zhangzhou tones can neutralize their contrasts across linguistic contexts.

As an extension to exploring the nature of Zhangzhou tones, this study adopts the eight-tone proposal. Figure 3 presents the normalized F0 pattern of Zhangzhou citation tones as a function of their corresponding normalized durations, which represents the central tendency of Zhangzhou as an independent variable. As indicated, the system included one rising (tone 1), three levels (tones 2, 5, and 8), one mid-low level with a final fall (tone 7), and three falling contours (tones 3, 4, and 6). Tones 4 and 6 both present a mid-high [41] falling trend. Similarly, tones 2 and 8 show a mid-low-level contour [22]. They are treated as contrastive primarily because of their considerable differences in other parameters, such as duration, voice quality, vowel quality, and syllable coda type [16-17].

4. Dominance of Zhangzhou Tone Sandhi

Zhangzhou presents a typical right-dominant tone sandhi system [16-17], [22]. Tones in the non-right-most positions, referred to as the sandhi context, alter their realizations at both phonetic and phonological levels, whereas tones in the rightmost position maintain their realization categorically similar to their citation forms. This property of right dominance can be straightforwardly demonstrated using scientifically objective patterns, as shown in Figure 4. It plots the normalized F0 contours of Zhangzhou Tone 3 across three linguistic contexts, which are derived from quantifying and normalizing 21 native speakers’ monosyllabic and disyllabic utterances.
Tone 3 presents a high-falling contour in citation but is changed to rising when it occurs in the phrase-initial position, whereas the falling contour shape is preserved in the phrase-final context. The two important properties of tonal alternation in the non-rightmost position and categorical preservation in the rightmost position justify the right dominance of Zhangzhou tone sandhi. However, tonal forms in the rightmost context are not always the same as their citation forms. They were observed to be phonetically sensitive to preceding tones and presented variations that could be statistically significantly different. This is discussed in detail in Section 6.

5. Syntactic Sensitivity of ZZ Tone Sandhi

The process of tone sandhi in Zhangzhou is not phonologically specified; instead, the sandhi domain is aligned with the boundary of syntactic phrases [16-17]. Within a given syntactic phrase XP, where X is categorically variable, ranging over nouns (N), verbs (V), prepositions (P), adjectives (Adj), and adverbs (Adv), among others, tones in the non-phrase-final position undergo categorical alternation while preserving the citation forms in the phrase-final position. This syntactic sensitivity is well demonstrated in Figure 5 (C=citation, S=sandhi). The sentence was elicited from a field recording with a female speaker, Yonghui Huang, with respect to her socio-background investigation in 2015.

$$(3)$$

Figure 5: Normalized F0 patterns of tone 3 across contexts

As indicated, the pitches of the tones at the phrase-final boundaries were largely the same as their corresponding citation values, whereas the pitches of the tones at the non-phrase-final position were categorically different from their citation forms. Thus, it is plausible to propose that the tone sandhi domain in Zhangzhou was syntactically sensitive. One exception can be seen on the monosyllabic morpheme /gwo3/ ‘1sg.’ whose pitch value is raised to an extra-high level, although it also marks the rightmost boundary. However, this exception is predictable and understandable from the articulatory perspective. Because it serves as the subject of a sentence, it is reasonable for speakers to raise their pitch register to introduce their own socio-linguistic background. Another special point is that the tone sandhi process in Zhangzhou is sensitive only to the boundary of syntactic phrases and is irrelevant to their internal structures. This is because tones behave consistently regardless of whether the related syntactic phrase is an NP, VP, AP, or AdvP. This property is illustrated in Table 1. All phrase-initial morphemes alter their pitch contour to mid-rise [35] from high-falling [51] in citation, regardless of their category: NP, VP, AP, or AdvP.

6. Phonological Inertness of ZZ Tone Sandhi

The tone sandhi process in Zhangzhou was phonologically inert. The parametric realizations of the tones at the sandhi (non-phrase-final) position were not affected by the categories of their subsequent tones within the same sandhi domain. In other words, regardless of their tone, individual sandhi tones are observed to have the same realizations. This property can be justified in Figure 6, which is derived from quantifying approximately 24, 696 (=588 tokens × 2 syllables × 21 speakers) F0 values. The left panel shows normalized F0 patterns of individual sandhi tones, while the right panel shows the pairwise t-test results on examining whether their F0 offsets are affected by their subsequent tones in a statistically significant way.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Citation</th>
<th>Sandhi</th>
<th>Example</th>
<th>Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>3+3</td>
<td>[51+51]</td>
<td>[35+51]</td>
<td>/ŋ3.&quot;swim/ ‘seawater’</td>
<td>NP</td>
</tr>
<tr>
<td>3+3</td>
<td>[51+51]</td>
<td>[35+51]</td>
<td>/s3.&quot;horses/ ‘wash horse’</td>
<td>VP</td>
</tr>
<tr>
<td>3+3</td>
<td>[51+51]</td>
<td>[35+51]</td>
<td>/n3.&quot;very well’</td>
<td>AP</td>
</tr>
<tr>
<td>3+3</td>
<td>[51+51]</td>
<td>[35+51]</td>
<td>/k3.&quot;kny/ ‘hurry up’</td>
<td>AdvP</td>
</tr>
</tbody>
</table>

Table 1. Examples of irrelevance to internal syntactic structures in Zhangzhou tone sandhi.

Figure 6: Normalized F0 patterns of individual phrase-initial tones across eight tonal combinations from 21 speakers.
As indicated, all phrase-initial tones presented a consistent F0 contour shape as a function of the following tones. For example, Tone 1 showed a mid-leveling contour across eight tonal combinations. This justifies their phonological inertness in the category of tone that occurs after them. However, three (tones 1, 2, and 3) out of eight phrase-initial tones have two phonetic variants, with their offset values slightly higher before tones 2, 7, and 8 than before the other tones. This reflects their phonetic sensitivity to F0 onsets of the following tones, which will be discussed in the next section.

7. Phonetic Sensitivity of ZZ Tone Sandhi

The Zhangzhou tone sandhi process is phonetically sensitive. This can be justified in Figure 6, in which three out of the eight phrase-initial tones have two variants in their F0 offset realizations. Figure 7 shows the normalized F0 patterns of individual phrase-final tones as functions of their preceding tones [22]. As can be seen, all tones except Tone 3 have two phonetic values in their onset positions. In general, F0 onset tends to be statistically significantly higher after tones ending in a [-low offset] and [-falling] but lower if the preceding tone ends in [+falling] and [+low offset] [22].

Phonetic sensitivity reflects the across-linguistic tendency of tonal coarticulation, particularly concerning the carryover effect (i.e., the influence of the preceding tones on phrase-final tones) [29-30]. However, this linguistic finding contradicts and challenges the conventional default principle for the identification of the right-dominance of tone sandhi in Sinitic languages, which considers that the right-dominant tones are directly related to their corresponding citation tones, with the citation values preserved and unchanged [1-2]. However, this does not fit in the phonetic reality in Zhangzhou Southern Min.

8. Morphological Nature of ZZ Tones

Zhangzhou tones can be posited as morphological in nature. Each lexical tone is treated as a single morpheme containing two or more alternative allomorphs (tonemes) that are all abstractly stored in the mental grammar of native speakers, but are phonetically distant from each other. This proposal was based on three main phonetic realities. (a) The sandhi (non-rightmost) tones are unrelated to their corresponding citation tones at both phonetic and phonological levels. This can be observed by comparing the normalized F0 contours in Figure 3 and 6. For example, tone 1 is realized as a rising contour in citation, but is changed to a mid-level in the sandhi (non-rightmost) position. Thus, the two tonal forms should be treated as independent systems. (b) The rightmost (non-sandhi) tones were categorically similar to their corresponding citation forms. For example, as plotted in Figures 3 and 7, tone 1 presents a rising contour in both citation and phrase-final contexts. Although most phrase-final tones vary in their onset height, they are predictable. Thus, phrase-final tones and citation tones should be treated as belonging to the same system. (c) Tonal neutralization occurs across linguistic contexts; determining which form is based and which form is derived appears difficult. This is illustrated in Figure 8. Tones 1 and 2 differ in citations but neutralize their contrast (rising vs. low level) to a mid-level contour in the sandhi context. Likewise, in Figure 9, tones 2 and 8 both present a low citation level, but they differ considerably in the sandhi context (mid-level vs. mid-falling). The direction of the tonal alternation is thus indetermined.

Figure 8. Tonal neutralization in the sandhi context.
Incorporating these three factors, it is thus appropriate to consider that the nature of the Zhangzhou tones is morphological, as schematized in Figure 10. In the context of general syntactic phrases, the relation between the sandhi (the non-phrase-final; non-rightmost) tones and citation tones are morphophonemic, while the relation between the non-sandhi (the phrase- final; rightmost) tones and citation tones is allophonic.

Figure 10. Morphological nature of Zhangzhou tones.

This morphological treatment is similar to the lexicon-based interpretation of allomorph selection in Taiwanese tone sandhi [13-15], but has different phonetic bases. For example, this study is grounded in examining acoustic data from 21 native speakers across three linguistic contexts in Zhangzhou Southern Min, whereas the Taiwanese proposal is based on production experiments [14-15] on citation tones or summarizing existing arguments from multiple sources [13].

10. Conclusion

As discussed, the encoding of tone sandhi in Zhangzhou Southern Min involves substantial interactions among linguistic levels comprising phonetics, phonology, syntax, and morphology. The exploration substantially stretches and advances our knowledge of tone sandhi as an important language phenomenon in Southern Chinese dialects. We hope to serve as a model to investigate tone sandhi in Sinitic languages and in other tonal languages where relevant.

11. References