






## ARTICLE

# The Interaction Between Tone and Duration in Du'an Zhuang

Jeremy Perkins<sup>1,\*</sup> , Seunghun J. Lee<sup>2</sup>  and Julián Villegas<sup>1</sup> 

<sup>1</sup>University of Aizu and <sup>2</sup>International Christian University & University of Venda

\*Corresponding author. Email: [jerperkins@gmail.com](mailto:jerperkins@gmail.com)

(Received 11 August 2023; revised 14 November 2024; accepted 3 December 2024;  
first published online 4 April 2025)

## Abstract

This research investigates the tone system of an understudied language, Du'an Zhuang and its interaction with duration. Cross-linguistically, tones tend to be less complex in shorter duration contexts. In Du'an Zhuang, syllable type provides these contexts: There are six contrastive tones among unchecked syllables with longer rhyme duration, but this is reduced to four tones in shorter duration checked syllables. Acoustic analyses of f<sub>0</sub> and duration from six native speakers were performed to check whether tonal complexity is reduced in the shorter duration checked syllables. The results showed this was true with some exceptions. The two tones in CVVO syllables corresponded to the two least complex tones; however, one of the two CVO tones included a more complex rising tone. This rising tone exhibited a reduced f<sub>0</sub> excursion though. Finally, there is a two-way phonological vowel length contrast in Du'an Zhuang, which necessarily interacts with syllable type via its effect on rhyme duration. However, based on our vowel duration measurements, this vowel length contrast only exists in unchecked syllables with sonorant codas, the only syllable type where rhyme duration and vowel duration could possibly differ. In this context, a sonorant coda contributes to the syllable's rhyme duration, but not to vowel duration, allowing syllable type and vowel length to contrast independently, only in this phonological context.

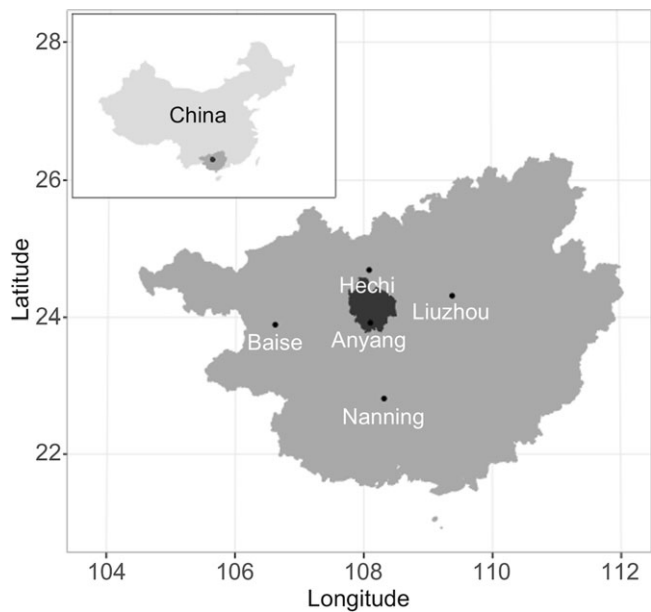
**Keywords:** tone; Zhuang; acoustic analysis; duration; contour tone

## 1. Introduction

The tone system of Du'an Zhuang is described in this research via phonetic analysis. In Chinese traditional descriptions, these have been based on qualitative empirical observations of tone. However, with access to new tools, we can probe further into previously reported tonal systems. This research is one such study, providing phonetic detail of the tone system of Du'an Zhuang.

Du'an Zhuang belongs to the Hongshui He (红水河) language group and is a variety of Northern Zhuang. Figure 1 situates the locations of the Du'an community within a map of Guangxi province, China. There are at least 16 million people who identify themselves as ethnically Zhuang, with about half a million people living in Du'an county (National Bureau of Statistics, 2012).

© The Author(s), 2025. Published by Cambridge University Press on behalf of The International Phonetic Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



**Figure 1.** The dark shaded region is the approximate region where Du'an Zhuang is currently spoken within Guangxi Province, China. Top left: The location of Guangxi province within China.

The tradition of describing Zhuang is based on the standard variety spoken in Wuming (Wei & Qin, 1980). Several studies on Zhuang dialects have been published (Qin, 1996; Zhang et al., 1999; Zhang & Qin, 1993), including a grammar study (Qin, 1995), in addition to an extensive dictionary (GZZW, 2005). There is also a recent description of Du'an Zhuang (Li, 2011) with no acoustic analysis. The inventories of consonants and vowels for Du'an Zhuang is shown in Tables 1 and 2 respectively and we give a brief overview of the phonology below.

**Table 1.** Consonant inventory of Du'an Zhuang (Zhang et al., 1999: 101)

<b>p</b>	<b>b</b>	<b>t</b>	<b>d</b>	<b>k</b>	<b>ʔ</b>
<b>m</b>	<b>n</b>	<b>ɲ</b>	<b>ŋ</b>		
<b>f</b>	<b>θ</b>	<b>ɸ</b>		<b>h</b>	
	<b>r</b>	<b>j</b>			
	<b>l</b>				

**Table 2.** Vowel inventory of Du'an Zhuang (based on Zhang et al., 1999: 102; arrangement ours)

<b>i</b>	<b>i:</b>	<b>ɯ</b>	<b>ɯ:</b>	<b>u</b>	<b>u:</b>	<b>ɨ</b>	<b>ɨ:</b>	<b>ɯ</b>	<b>ɯ:</b>
<b>e</b>	<b>e:</b>	<b>ə</b>	<b>o</b>	<b>o:</b>	<b>ei</b>	<b>o:i</b>	<b>e:u</b>	<b>əu</b>	<b>ou</b>
<b>a</b>	<b>a:</b>		<b>ai</b>	<b>ai:</b>			<b>au</b>	<b>au:</b>	<b>au:</b>

**Table 3.** Previous descriptions of the tonal systems of Du'an Zhuang (Zhang et al., 1999; Castro & Hansen, 2010; Li, 2011)

	Tone	Description	Example Word
Unchecked Tones	1	High falling	na
	2	Low rising-falling	na
	3	High rising-falling	na
	4	Low rising	na
	5	Mid level	da
	6	Mid or low falling	da
Checked Tones	7	High level	tap
	8	Mid level	mak
	9	Low rising	ta:p
	10	Low falling	ma:t

All consonants can occur in onset position but only some of these also occur in coda position (shown in bold font in Table 1). [b] and [d] are phonetically pre-glottalized [ʔb] and [ʔd] and [r] occurs with frication according to Zhang et al. (1999). Clusters [kj], [kw], and [ɲw] are permitted in onset position. There is a binary vowel length contrast in many, but not all rhymes and not all possible VC sequences are attested either: short [e] is absent, except in syllables with [k] codas; [u] does not occur in syllables with labial codas [m, p]; and [ə] is only attested in open syllables. Diphthongs are only attested in open syllables, never in closed syllables, and could be analyzed as rhymes with glide codas, if not for the apparent contrast between [au] and [au̯].

Previous descriptions of the tonal system of Du'an Zhuang state that it has six tones that occur in unchecked syllables and four tones that occur in checked syllables, defined based on f<sub>0</sub> contour contrasts (Zhang et al., 1999; Castro & Hansen, 2010: Table 3, tonal region A; Li, 2011). 'Checked' refers to syllable rhymes with obstruent codas and 'unchecked' refers to open syllables and those with sonorant codas. Previous descriptions of the tonal system are summarized in Table 3, with a more detailed summary included in Appendix A. We adopt the numbering system using tones 1 to 10 used in previous traditional Chinese field work for expository convenience, but not as a claim of tonal identity. Both Zhang et al. (1999: 101) and Li (2011: 18) separately described each of the ten tones, noting that the four checked tones have the same f<sub>0</sub> contours as four of the unchecked tones (3 & 7, 4 & 9, 5 & 8, 6 & 10), a common situation in Tai-Kadai languages (Pittayaporn, 2009). This implies a six-tone system in unchecked syllables, which is reduced to four in checked syllables. The question of whether allotonic correspondences exist between the four tone pairs above is addressed in Sections 3 and 4. Notably, neither of the previous studies offered phonetic evidence, relying on impressionistic observations of the tonal system instead. As a result, a quantitative phonetic study of the tonal system of Du'an Zhuang is still needed.

This research analyzes the tonal system of Du'an Zhuang and its interaction with rhyme duration based on phonetic measurements. Lexical tone is expressed cross-linguistically via pitch (f<sub>0</sub>), rhyme duration and phonation. As can be seen above, syllable type plays an important role in the Du'an Zhuang tonal system since it apparently splits the tonal system into separate systems for unchecked syllables and checked syllables. Rhyme duration is investigated here as the most likely phonetic expression of syllable type. At the same

time, Du'an Zhuang also supports a phonological vowel length contrast between short and long vowels. Vowel duration is investigated here as the most likely phonetic correlate of vowel length. Vowel length and syllable type are independent phonological categories, but they are necessarily phonetically linked to one another: Syllable rhymes are composed of a vowel and optionally a following coda consonant. As such, rhyme duration and vowel duration are necessarily correlated, and are equivalent to each other in open syllables and in closed syllables with obstruent coda consonants. The only context where they can differ is in syllables with sonorant coda consonants.

Interestingly, Du'an Zhuang's vowel length contrast is limited to exactly that context: syllables with sonorant coda consonants. Otherwise, vowel length is redundantly assigned: All open syllables have long vowels and in checked syllables, vowel length piggybacks with tone: Tones 7 and 8 always have short vowels and tones 9 and 10 always have long vowels. Functionally, this means that the vowel length contrast exists in the only possible context where it can be phonetically distinguished from syllable type: A sonorant coda contributes its duration to the syllable's rhyme duration, but not to vowel duration, allowing syllable type and vowel length to exist independently in this context.

As mentioned, syllable type is expected to play an important role in the Du'an Zhuang tonal system since different sets of tones exist in unchecked and checked syllables. The most likely phonetic correlate, rhyme duration, is then predicted to differ along that dimension: Unchecked syllables should have longer rhyme duration than checked syllables. However, phonological vowel length can also take on this role and in many languages it does so (i.e. Somali and Navajo (Zhang, 2001, 2004)), and in those cases, vowel duration is predicted to correlate with tones. In fact, it is even possible that both vowel length and syllable type may be involved. Du'an Zhuang's checked tone system redundantly assigns tone based on a vowel length difference as explained above. Therefore, the first goal of this investigation is to determine the roles of syllable type (via phonetic rhyme duration) and phonological vowel length (via phonetic vowel duration) on the tones in Du'an Zhuang.

Once the role of duration in the tonal system is determined, Du'an Zhuang can then be assessed in terms of the cross-linguistic typology of tone and duration. Duration and tone interact with each other both phonetically and phonologically. The autosegmental status of tone allows it some degree of independence from segmental phonology. However, there is evidence from languages like Thai and Zahao where rising tone is banned in syllables with obstruent codas, that laryngeal features interact with tone. Yip (1982) and Morén & Zsiga (2006) assume that there is some laryngeal constriction in these obstruent codas, and that tone and laryngeal features interact phonologically by virtue of a shared laryngeal feature node. Regarding duration, Yip (1982) notes that in Zahao, syllables with rising tone are somewhat longer than those with level low or high tones but argues that this is a 'late phonetic fact', rather than a phonological one.

However, this kind of interaction between duration and contour tone can be seen as a phonological fact when considering other languages. Duanmu (1990) states that tone-bearing ability is directly related to rhyme length, such that longer rhymes correlate with greater tone-bearing ability. The concept of tone-bearing ability was elucidated further by Zhang (2001, 2004), who showed that the possible degree of complexity of a tone tends to be higher in certain contexts where phonetic duration is longer, and likewise tonal complexity tends to be lower in contexts where phonetic duration is shorter. Zhang defines tonal complexity across three particular dimensions that are relevant for duration: (1) Tones with more pitch targets are more complex; (2) Tones with greater pitch excursions are more complex; and (3) Rising tones are more complex than falling tones, all else being equal. A tone that is more complex requires a longer duration then. There is cross-linguistic evidence for these generalizations, coming from empirical, functional, articulatory and perceptual realms. For example, rising tones require longer durations to produce than



falling tones (Ohala & Ewan, 1973), and so rising tones are argued to be more complex than falling tones by Zhang (2001, 2004).

The empirical evidence comes from a wide range of languages. First, the distribution of contour tones is often limited to more prominent positions that are afforded longer sonorous durations phonetically (Gordon, & Ladefoged, 2001; Zhang, 2001; Yu, 2003). This kind of restriction is seen in Somali and Navajo, where contour tones are restricted to long vowels; in Xhosa, where contour tones are restricted to stressed syllables; and in Thai, where rising tones are restricted to unchecked syllables (Abramson, 1962; Zhang, 2004).

On the perception side, contour tones are perceived as longer than level tones (Lehiste, 1976). Among level tones, high level tones are generally perceived as longer than low tones (Yu, 2010). However, despite this, low tones tend to have longer rhyme durations than high tones cross-linguistically (Gandour, 1977). Perhaps because of this lack of alignment between perception and production among level tones, there are very few cases of phonological processes where duration interacts with level tones. Instead, most phonological interactions between tone and duration involve contour tones, where multiple factors prefer longer sonorous rhyme durations.

However, languages may differ on which particular factor is referenced in the phonology (vowel length, stress, syllable type, etc.), and to what extent phonetic alterations are made to reduce tonal complexity. They may reduce the  $f_0$ -excursion amount or lengthen the duration to accommodate the tone. For example, in Hausa, falling tones can occur on CVV, CVR and CVO, but not CV syllables. However, falling tones in CVO syllables occur with lengthening of the vowel (Gordon, 1998) and with significant reduction of the  $f_0$  excursion (Zhang, 2001). Thus, both  $f_0$  and duration of falling tones vary in these different syllable types, creating two different falling allotones in Hausa. This specific example illustrates how allophonic correspondences of tones can arise, in this case with syllable type providing the context for the variation.

Du'an Zhuang's tonal system provides an additional test case for such allotonic correspondences involving contour tones. Recall that the four checked tones were noted to have identical tonal contours to four of the unchecked tones in previous research, suggesting such allotonic correspondences. In particular, it is predicted that contour tones in checked syllables may be simplified relative to their allotones in unchecked syllables, assuming that checked syllables have shorter rhyme durations. However, unlike Hausa, where only open syllables can have a vowel length contrast, Du'an Zhuang includes a vowel length contrast in unchecked syllables with sonorant codas (CVVN and CVN). Recall that in checked syllables, vowel length is redundant depending on the tone: two tones occur with a long vowel (CVVO) and two other tones always occur with a short vowel (CVO). This then provides an additional possible context (vowel length) that could modulate the allotonic correspondence. Phonologically short vowels are likely to correspond to phonetically shorter vowel durations, which in turn may result in shorter rhyme durations. One major aim of this research then is to explore how syllable type and vowel length modulate allotonic correspondences in Du'an Zhuang.

To accomplish this, it is necessary to consider methods to compare tones across contexts. One way is to limit these comparisons to allotones, but this requires a method to identify allotones in the first place. In this investigation, rather than seeking to determine whether allotonic correspondences exist, tones are chosen for comparison based on similarity of relevant phonetic properties, here  $f_0$  contour. It should be noted that definite conclusions on the presence of allotones would require more evidence, including morpho-phonological alternations, evidence of historical changes, or results of perceptual experiments. Regarding historical evidence, Pittayaporn (2009) notes that there is often tonal identity in modern Tai-Kadai languages between particular checked and unchecked tones. In fact, Zhang et al. (1999: 24) gives historical correspondences between each Proto-Tai tone category and each modern Zhuang tone. In other words, checked and unchecked

tones are often identical in Tai-Kadai languages. However, large duration differences make this identity less clear. If the duration of a given tone is reduced in checked syllables, then either the phonetic contour target is maintained but compressed over a shorter duration or it is truncated to reach a particular part of the contour target. In either case, some phonetic modification is necessary. Therefore, rather than referring to ‘tonal identity’, we use the term ‘allotone’ to describe these two phonetic reflexes of the same tone category. Thus, in our research, allotonic correspondences are seen as more likely in cases where a strong match is found with respect to  $f_0$  contour. Beyond this, it is not relevant for these purposes whether the two tones are in fact allotones – it is enough that they are similar and can thus be compared in the context of tonal complexity and its interaction with duration as described above.

The range of phonetic variation that may be expected in allotones in environments with shortened duration must also be considered, as mentioned briefly above: One possibility is that the  $f_0$  excursion of a contour tone will be reduced, while maintaining the same basic contour shape. A second possibility is that part of the  $f_0$  contour will be truncated, so that an edge of the tone is cut off. In either case, the visual presentation of the tones is crucial in correct identification of similarities. In the first case, where  $f_0$  contour is preserved with reduced excursion, to identify visual similarity in  $f_0$  contour shape,  $f_0$  should be plotted against normalized time in both environments. However, in the second case, where  $f_0$  contours are truncated, normalized  $f_0$  plots would result in incorrect conclusions. Instead,  $f_0$  should be plotted against time without any normalization, so that the partial match of the  $f_0$  contours would allow identification of allotones. There is also a third possibility – that the allotone in the environment with shortened duration would involve some relative lengthening, like Hausa, in which case, normalized time plots would again be appropriate.

These three possibilities are not mutually exclusive – languages may employ any subset of them. For example, Hausa combined the first and third strategies and so a proper investigation should consider all of them. Of course, one final possibility is that the tone will surface identically, despite the shortened duration. This is predicted to be more likely for less complex tones, such as level tones and in environments with more moderately shortened rhyme duration.

Next, the possible environments where duration may be shortened need to be considered. As mentioned, in Du’an Zhuang, these involve a combination of vowel length and syllable type. This situation arises in other languages as well, such as Thai, Cantonese and Navajo. Zhang (2004) noted that Thai and Cantonese both have contrastive vowel length, but that contour tones are restricted only in checked syllables, without referring to vowel length. In Thai, rhyme duration and not vowel duration correlates with the tonal split, but in Cantonese, while rhyme duration dominates, vowel duration also plays a role. However, in Navajo, the opposite is true: It is vowel length, and not syllable type that is the determining factor for the distribution of contour tones.

Therefore, one possibility is that Du’an Zhuang may resemble Thai and Cantonese rather than Navajo. This would follow from the descriptive account, that syllable type, and not vowel length, is the dimension that is most relevant for tonal distribution in Du’an Zhuang as shown in Table 4. According to Zhang’s theory, the same prediction for Thai would then apply to Du’an Zhuang (Zhang, 2001: 155): ‘[Un]checked syllables should have longer sonorous rhyme duration than checked syllables. In particular, CV > CVVO, CVN > CVVO’. In other words, the fact that a syllable is unchecked may trump the effect of vowel length on rhyme duration in Du’an Zhuang.

However, this is not the only possibility – vowel length may still be relevant in addition to syllable type. In Cantonese, both vowel duration and rhyme duration together were active. This possibility is highlighted in the descriptions of Du’an Zhuang’s tonal system in checked syllables, where a pair of level tones (tones 7 and 8, high level and mid level, respectively)

**Table 4.** Predicted hierarchy of rhyme duration differences via equal contributions from syllable type and vowel length

Predicted Rhyme Duration		Syllable Type (Unchecked > Checked)	Vowel Length (Long > Short)
Longest duration	CVV, CVVN	✓	✓
Medium duration	CVN or	✓	×
	CVO	×	✓
Shortest duration	CVO	×	×

Checkmarks correspond to cells predicted to have longer rhyme duration given their phonological status in each column.

were identified that are found in CVO syllables and a pair of contour tones (tones 9 and 10, low rising and low falling respectively) that are found in CVVO syllables. This difference involves reference to vowel length and not syllable type, and so following Zhang's logic, and assuming the descriptions are correct, we may expect both vowel length and rhyme duration to be relevant, both providing contexts for shortened duration affecting allotones. As a simple illustration, in the case that both factors contributed equally, we would expect rhyme duration to essentially have three levels, as outlined in Table 4.

The comparison between CVN and CVVO syllables is crucial, where the two factors, syllable type and vowel length, are pitted directly against each other. In Du'an Zhuang, both vowel length and syllable type contrasts coexist, but their phonetic expressions (rhyme duration and vowel duration) are inherently co-dependent. Lengthening a vowel also lengthens a rhyme, all else being equal. However, in Thai, for example, all else is not equal: Nasal codas are lengthened in CVN syllables such that the rhyme duration actually exceeds that of CVVO syllables (Zhang, 2004; Morén & Zsiga, 2006). In this way, syllable type contrasts are expressed via rhyme duration, and this does not interfere with the vowel length contrast, which is expressed separately via vowel duration. This is a likely configuration in any language with interactions between tone, contrastive syllable type and vowel length, like Du'an Zhuang. By measuring rhyme duration for CVN and CVVO syllables in Du'an Zhuang, this will confirm whether and to what extent syllable type and vowel length are relevant for tonal contrasts.

While our research is mostly focused on syllable type and vowel length, duration differences inherent to tones have also been widely observed in tonal languages, and sometimes are used as cues in tonal perception. Recent phonetic studies have shown duration effects for distinguishing lexical tones for Mandarin Chinese (Fu & Zeng, 2000; Liu & Samuel, 2004), White Hmong (production: Esposito, 2012; perception: Garellek et al., 2013), Takhian Thong Chong (DiCanio, 2009), and Sgaw Karen (Brunelle & Finkeldey, 2011). As a result, lexical tone itself may influence rhyme duration, and so effects of tone identity on rhyme duration need to be accounted for as well.

The aims of this research are thus twofold. First, we explore whether and to what extent rhyme duration is affected by syllable type and vowel length. A sub-goal here will be to analyze vowel duration along the same factors. This will provide some insight on how vowel length and syllable type differences can co-exist as independent phonological dimensions despite their phonetic co-dependence. Second, once the factors affecting rhyme duration have been identified, Du'an Zhuang's tone system can be assessed in the context of the cross-linguistic typology for the interaction between contour tones and duration. It is predicted that the tone system in short-duration contexts (checked syllables or short vowels) should not be more tonally complex than the tones found in contexts with longer duration

(unchecked syllables or long vowels). By measuring  $f_0$  over time along with rhyme duration, this prediction can be confirmed by comparing these phonetic properties from tones in each context. Finally, given these  $f_0$  and duration results, we check for specific allotonic correspondences based on phonetic similarity (of  $f_0$  and relative rhyme duration) between tones occurring in longer duration and shorter duration contexts.

In addition to  $f_0$  and duration, phonation is often relevant for tone. In Perkins *et al.* (2016), a single speaker of Du'an Zhuang produced tones 4 and 6 with creaky phonation, suggesting that creakiness may play a role in these two tones. This possibility was checked via measurements of spectral tilt (Gordon & Ladefoged, 2001) and psychoacoustic roughness (Villegas *et al.*, 2020) for the six Du'an Zhuang speakers who participated in this study, but none of them consistently used creaky phonation in any of the unchecked tones. However, in checked tones, evidence of creaky phonation (via increased roughness) was seen in the latter part of the vowel adjacent to the obstruent coda, beginning as early as 50% and as late as 80%, depending on the speaker. This association between checked syllables and creaky phonation has been attested in other languages, include some Wu dialects of Chinese (Shen, 2010). This is likely a phonetic effect of the glottal gesture associated with the coda, and there was no evidence of creakiness correlating with any particular tone. As a result, an analysis of creaky phonation is not presented here. It is possible that some speakers do use creaky phonation as part of the tonal system though, and that a larger sample of Du'an Zhuang speakers may confirm this. Even in languages where  $f_0$  is the primary cue, phonation is sometimes relevant as a secondary cue. For example, an anonymous reviewer noted that in related languages, proto-tone 4 (C2 category) sometimes exhibits creaky phonation. Furthermore, in Mandarin Chinese, creaky phonation commonly occurs in tone 3 and sometimes tone 1, in production of very low or high pitch targets (Kuang, 2017).

Finally, establishing a tone system in a variety where there is another well-documented lect, Wuming Zhuang in this case, offers several challenges to researchers. Dictionaries and other published materials usually offer word lists in the well-documented lect. While some dialects are very similar, differences should be expected, and all items need to be checked after data collection. In the case of Du'an Zhuang, the differences with the standard Wuming lect are substantial and dialects of Zhuang reportedly show lexical and tonal variations so that some dialects are mutually unintelligible (Castro & Hansen, 2010; Kullavanijaya, 2001). Balanced word list creation, controlling for effects of vowel quality, consonant context, etc. becomes difficult as well, since these dimensions may vary from what is listed in sources, yielding utterances that often differ from what was expected. A lack of literacy in Zhuang also creates problems in creating balanced word lists since stimuli must be shown via a non-native script, and thus the intended lexical item has a lower probability of being offered. Systematic approaches to these problems have been offered in previous research; for example, Coupe (2014) outlines advice and methods useful in documenting tone languages with no previous known tonal analysis. The following section details the methods adopted here for Du'an Zhuang.

## 2. Methods

### 2.1 Consultants

Six speakers (four females and two males) of Du'an Zhuang were recorded. All speakers grew up in a Du'an town and were between the age of 20 and 23. They reported speaking Zhuang with family, friends, and classmates, but they also have spoken the Guangxi variety of Putonghua Chinese from a young age. In addition to Putonghua, most speakers spoke some English; three speakers reported that they also spoke the Guiliu dialect. None of the speakers reported any hearing problems that affected their speech. Finally, as Du'an Zhuang

is a minority language in China, it is difficult to find large numbers of speakers. As a result, with only six speakers, this study is very limited in its ability to generalize across the Du'an Zhuang population and future work with a larger pool of speakers is needed to confirm the findings here.

## 2.2 Vocabulary

One-hundred-and-forty-three monosyllabic Zhuang words were produced based on 133 Chinese orthographic prompts from Qin (1996) (see Appendix B for a complete word list). Chinese orthography was used because participants were not familiar with the romanized Zhuang script and were all comfortable with the Chinese characters. Because words were not elicited in Zhuang script, participants had to choose a Zhuang word matching the meaning of the Chinese prompt. For a single prompt, it was possible that more than one Zhuang word could be elicited. As a result, there were a large number of words produced that did not match the intended Zhuang word from the dictionary. These words were excluded from the analysis.

Words were selected from the dictionary that contained alveolar, palatal, or velar obstruent onset consonants. Both long and short vowels of all vowel qualities were used in words containing codas; however, open syllables only contain long vowels in Zhuang, so there are no words with open syllables and short vowels. Words with unchecked tones were chosen with a nasal coda at any place of articulation (CVN, CVVN) or without any coda (CVV). Lexical frequency differences among the tones existed. For example, there were not many words with tones 4, 8, 9, or 10 in Qin (1996), resulting in those tones being under-represented.

Table 5 shows the distribution of the words by tone, with each word repeated five times. Note that some Chinese orthographic prompts in our word list had more than one corresponding Zhuang word listed in the dictionary, with speakers sometimes differing on which word they produced. Because of this, a given prompt led to the production of multiple possible Zhuang words, and these words often had different tones. This, in turn, resulted in the number of possible phonologically unique Zhuang words produced (143) being greater

**Table 5.** Word list distribution by tone

	<i>Tone</i>	<i>Words Included</i>	<i>Tokens Analyzed</i>
Unchecked Tones	1	18	399
	2	14	265
	3	17	423
	4	7	177
	5	11	251
	6	17	355
Checked Tones	7	14	265
	8	8	158
	9	10	213
	10	6	141
	<i>All</i>	122	2647

than the total number of prompts (133) in the word list. For some words, all tokens were excluded however, resulting in a total of 122 unique words produced.

### 2.3 Procedure

Words in the list were read in isolation five times, resulting in 3,990 total prompts (133 words  $\times$  5 repetitions  $\times$  6 speakers). The words were presented in slideshows, one word per slide, with the corresponding Chinese orthography (see Appendix B for a complete word list). As mentioned in Section 2.2, there were many words that did not match the dictionary pronunciation with respect to tonal specification or segmental content. These non-matching words were transcribed with tone included, by inspection of f0 contour plots, requiring independent consensus on the tone from all three authors, and are listed in Appendix B. However, all words that differed from the dictionary were excluded from the analysis to avoid errors that may arise from the authors' tone judgments. In cases where participants were not familiar with a Zhuang word corresponding to a given Chinese prompt, participants were instructed to skip that item.

The participants were recorded in two adjacent, independently sound-attenuated rooms at a regional University in Guangxi, China. For all but two participants, Putonghua-speaking assistants aided with the recording procedure. The stimuli were displayed via a laptop with a 13-inch screen. The experimenters advanced the slides and verbally noted when participants offered a different form in reciting a given word. They also asked for repetitions of tokens that were unclear. The participants were given four practice items prior to the recording, and they took three breaks. They wore a head mounted unidirectional DPA 4088 microphone, connected to a Marantz PMD661 MK2 solid-state recorder with an XLR cable. The collected samples were stored in mono WAV files at a sampling rate of 48 kHz.

### 2.4 Annotation

The audio files were annotated using Praat (Boersma & Weenink, 2018). The syllable rhyme was marked to enable measurements of f0 and duration over the rhyme. In cases where a nasal coda was present, the vowel and nasal were separately annotated, with rhyme duration taken as the sum of the two durations. Obstruent codas were not annotated since they lacked any audible release. The decision to focus on syllable rhymes follows since nuclei and codas are the possible sites for tone (Yip, 2002) and since syllable rhyme duration has been found to interact with tonal systems as outlined in Section 1. The initial rhyme boundary corresponded with the first appearance of the vowel formants. The final rhyme boundary marked the point when voicing, F1, and F2 were no longer visually apparent. Rhyme boundaries were automatically moved to the nearest zero-crossing sample. Boundaries between vowels and nasal coda consonants were determined by spectral shifts and intensity decreases associated with a nasal consonant.

### 2.5 Acoustic analysis

The first research question asked whether and to what extent syllable type and vowel length separately affected rhyme duration. To address this question, rhyme duration was measured for all tokens via Praat. For checked tones, the obstruent coda consonants were not released, so only the duration of the vowel was measured. On the other hand, for unchecked tones with sonorant codas, the sonorant portion was included in the rhyme duration measurement. A linear mixed model was used to investigate rhyme duration differences. These were performed via lme4 (Bates *et al.*, 2015). The fixed effects of syllable



type (checked vs. unchecked), vowel length and tone were considered. Random intercepts were included for speaker, word item, and vowel quality. Type II ANOVAs using Wald chi-squared tests (for models without interaction effects) or Type III ANOVAs using Satterthwaite's method (for models with interaction effects) were performed to test for significance of fixed effects in each model using the *lmerTest* package in R (Kuznetsova, Brockhoff, & Christensen, 2017). Likelihood ratio tests using the maximum likelihood method were used to test for model significance, with an optimal model selected based on the results of these tests. The likelihood ratio was used to justify whether each fixed effect should be kept in the model. When appropriate, post-hoc Tukey tests for rhyme duration were performed to assess differences between each level for each fixed effect using the *emmeans* library in R (Lenth, 2023).

The second research question involved investigating the *f*<sub>0</sub> contours and rhyme durations over time. The Straight algorithm (Kawahara, Masuda-Katsuse, & De Cheveigné, 1999), as implemented in VoiceSauce (Shue et al., 2011), was used to extract *f*<sub>0</sub> measurements at 10 ms intervals over the syllable rhyme. Parameter settings used in Straight were 500 Hz for maximum *f*<sub>0</sub> and 40 Hz for minimum *f*<sub>0</sub> with maximum duration set to 10 s, which is appropriate for both male and female voices (Tsanas et al., 2014). *F*<sub>0</sub> normalization was performed to account for higher- and lower-pitched voices by dividing each measurement by the median *f*<sub>0</sub> across all utterances of a speaker, and then converting to cents, (one cent is equal to one hundredth of a semitone).

The prediction that tones in shorter duration contexts should be less complex than tones in longer duration contexts was tested visually via plots of *f*<sub>0</sub> over time. All plots were smooth scatterplots using *geom\_smooth* with the *gam* method in R's *ggplot2* library (Wickham, 2016; R Core Team, 2019). It was necessary to check two kinds of plots. This was the case because allotones may involve (1) preservation of complete contours, in which case time-normalized plots would be needed; or they may involve (2) part of a contour being cut off, in which case plots against time without any normalization would be needed.

To investigate the *f*<sub>0</sub> contours of the tone system across speakers, a generalized additive mixed model (GAMM) was created using the *mgcv* (Wood, 2017) and *itsadug* libraries (van Rij et al., 2023) in R. Three separate GAMM models were performed for unchecked, checked CVO, and checked CVVO syllables, due to the substantial differences in rhyme duration between these three syllable types. The dependent variable was *f*<sub>0</sub> in cents with parametric effects and smooths over (non-normalized) time for tone. Random smooths over time that varied by tone were included for speaker, vowel quality, word and onset type. An AR1 autoregressive error model was included to reduce the autocorrelation in the residuals (Sóskuthy, 2017). The number of knots, *k*, was set to 30 for all smooth terms in the unchecked model but set to the default, 10, for the two checked tone models. The constant, *m* was set to 1 for smoothing penalty estimation in the random smooth terms as recommended in Baayen et al. (2018) and Sóskuthy (2017). The dataset and R code used in the analysis is available at <https://github.com/jerperkins/DZData>.

## 2.6 Data exclusion

Of the 3,990 prompts presented to speakers, about a third (*n* = 1343) were not included in the final data set. 202 tokens (5.1% of prompts) were skipped by participants. A manual check of *f*<sub>0</sub>-contours was performed independently by the three authors to exclude items with incorrect *f*<sub>0</sub>-contours. 45 tokens (1.1% of prompts) were excluded as outliers compared with other repetitions of the same word by the same speaker, possibly due to production errors or to errors with the Straight *f*<sub>0</sub>-measurement algorithm (Kawahara, 2006). 69 tokens (1.7% of prompts) were excluded because of inconsistencies among repetitions of a single word over the course of the experiment. If at least two (out of five) such

tokens were discovered, then all five tokens produced by the speaker were excluded on these grounds. Consensus from independent checks by all three authors were required for any such exclusions.

In addition, two native Putonghua speakers checked the elicitations against the Chinese orthography, flagging any that they recognized as Putonghua. This review was done to check whether and to what extent interference with the speakers' Putonghua competency may have played a role. Only six such tokens were identified, suggesting that Putonghua competency did not interfere with the task to produce Du'an Zhuang words. These words were already excluded from the data set anyhow because they did not match Du'an Zhuang dictionary forms expected for the prompts.

An additional 56 words (1.4%) were excluded when speakers gave more than the prescribed five tokens for a given word. This could happen because of the one-to-many correspondence of Chinese prompts to Zhuang words: In cases where more than one prompt yielded the same Zhuang word for a given speaker, five tokens were chosen from a single prompt when possible. If more than one such possible prompt existed, then the prompt with the higher number of total tokens in the final data set across all speakers was selected.

In total, there were 851 words (21.3% of prompts) excluded because they did not match the expected Zhuang dictionary pronunciations. The majority of these cases involved segmental differences in the word. However, there were some cases where the segmental content matched the dictionary pronunciation but the wrong tone was used. To detect these cases, scatter plots of  $f_0$  over time for each token were visually checked by all three authors for consistency against words with the same tone for each speaker. In cases where a consensus was reached that the  $f_0$  contour of a word did not match other words with the same tone for that speaker, all tokens of the word were excluded from analysis due to dictionary mismatch. These cases were transcribed and listed in Appendix B. They may have resulted because of errors in the dictionary, historical tone changes, or from our speakers offering alternative lexical items for the prompt. The tone review was based on  $f_0$  alone in unchecked syllables because differences in rhyme duration among the unchecked tones were small. Duration was used together with  $f_0$  when reviewing tones in checked syllables, since the phonological vowel length difference between CVO tones 7 and 8 and CVVO tones 9 and 10 was diagnostic (see Section 3). The final data set retained about two-thirds of the word list for analysis (2,647 tokens).

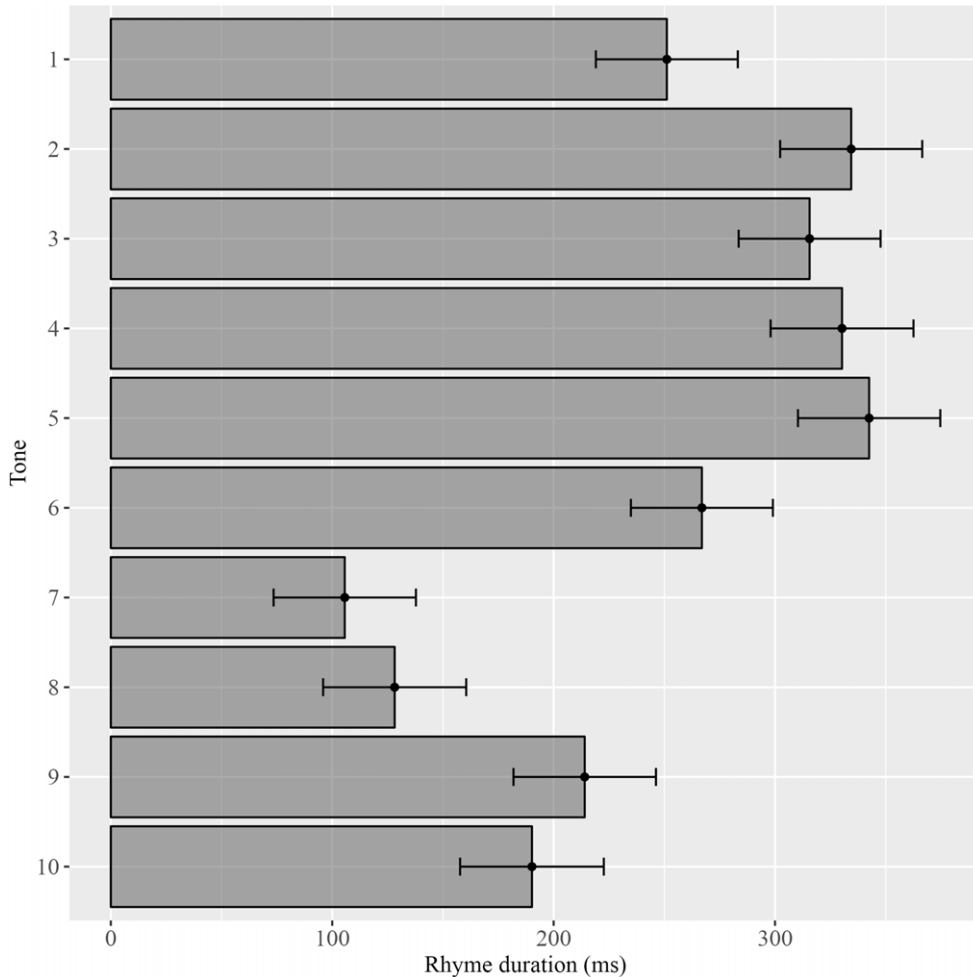
### 3. Results

#### 3.1 Duration

The first research question was to determine whether Du'an Zhuang's tone system is affected by syllable type (via rhyme duration) or vowel length (via vowel duration). The fact that the six-way tonal contrast is reduced to four tones among checked syllables suggests that syllable type is relevant. If this is the case, then rhyme duration should depend on syllable type, to a greater degree than vowel length. To address this, a linear mixed model was performed with rhyme duration as the dependent variable and syllable type, phonological vowel length and tone as independent factors. Random intercepts were included for speaker, word, and vowel quality.

This prediction was true: The effect of syllable type on rhyme duration was much larger than the effect of vowel length even though both were significant. In addition, a significant effect for tone ( $\chi^2(8) = 844.2$ ;  $p < .001$ ) was also discovered in a Type-II ANOVA. A likelihood ratio test showed that the effect of the interaction between tone and vowel length was not significant ( $\chi^2(5) = 7.7$ ;  $p = .170$ ), and so it was excluded from the final model. The coefficient of determination for the final model was  $R^2 = .831$ .





**Figure 2.** EMMs of rhyme duration (ms) by tone with 95% confidence intervals.

Pairwise post-hoc Tukey tests were performed via the *emmeans* package in R (Lenth, 2023). Estimated degrees-of-freedom used the Kenward-Roger method. All comparisons involve estimated marginal means of rhyme duration based on the model (EMMs). An EMM is a model-internal estimate of a response variable for each level of the variable. Unchecked syllables had significantly longer rhyme duration (EMM = 307 ms) than checked syllables (160 ms) ( $t(109) = 48.1$ ;  $p < .001$ ). Syllables with long vowels had significantly longer rhyme duration (238 ms) than those with short vowels (228 ms), but only by 10.3 ms ( $t(116) = 2.77$ ;  $p = .0065$ ). Rhyme duration also differed by tone with the following rank order: (5, 2) > 3 > (6, 1) > 9 > 10 > 8 > 7, with EMMs and 95% confidence intervals displayed visually in Figure 2. Rhyme duration of tone 4 did not differ from tones 2, 3, and 5, but was significantly longer than all other tones. Tones 1 and 6 were significantly shorter than other unchecked tones, raising the possibility that rhyme duration may be a secondary cue for these two tones; however, confirmation of this would require further investigation involving perceptual tests.

As mentioned, the effect size of syllable type on rhyme duration was more than ten times larger (147 ms) than that of vowel length (10.3 ms). This situation matches the

situation seen for Thai, where the effect of rhyme duration trumped the effect of vowel length. Therefore, in Du'an Zhuang, rhyme duration is modulated by syllable type and not vowel length. Like Thai, Du'an Zhuang has contrastive vowel length though, and so vowel duration, rather than rhyme duration is investigated as a correlate for vowel length.

To test this, a second linear mixed model was performed for vowel duration with the same fixed and random effect structure as the model for rhyme duration, but this time with Repetition included as a random intercept. There were 1,863 words included in the analysis (543 CVVN, 543 CVN, 354 CVVO, and 423 CVO). The 784 CVV words were not included because there is no vowel length contrast in open syllables.

The results showed that the effect size for phonological vowel length on vowel duration was much larger than syllable type this time. An additional significant effect was discovered for tone ( $\chi^2(8) = 53.1; p < .001$ ). The coefficient of determination for the final model was  $R^2 = .815$ . Pairwise post-hoc Tukey tests were performed for vowel duration for each level of syllable type, vowel length and tone. Estimated degrees-of-freedom used the Kenward-Roger method. As mentioned, long vowels had much longer vowel duration (208.8 ms) than short vowels this time (99.1 ms) ( $t(72.6) = 27.9; p < .001$ ). On the other hand, while significant, the difference between vowel durations in checked (160 ms) and unchecked syllables (148 ms) was about a tenth as large ( $+11.1$  ms;  $t(70.8) = 4.00; p < .001$ ).

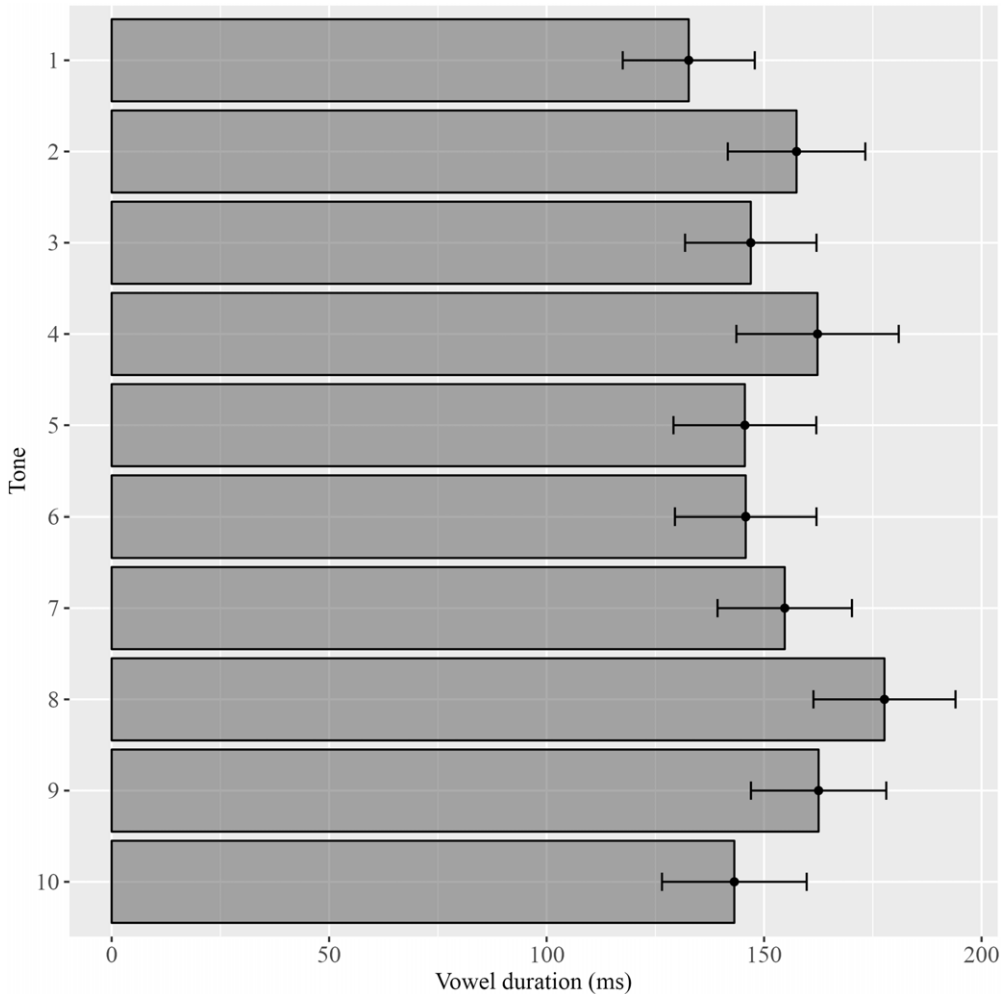
Among tones, tone 8 had a longer EMM for vowel duration than tones 1, 2, 3, 5, 6, 7 and 10 and tone 1 had a shorter EMM for vowel duration than tones 2, 4, 7, 8, and 9. The EMMs for vowel duration by tone are illustrated in Figure 3.

Notably, since vowel length is separately accounted for in the model, the EMM allows isolation of the duration effect due to each tone, separately from the effect of vowel length. Thus, tone 8 (only CVO) is associated with vowel duration lengthening, even though the raw duration is very short. The same is true in the other direction, where tone 10 (only CVVO) has longer vowel duration but once the effect of vowel length is removed, the inherent vowel duration for tone 10 is relatively shorter than tone 8.

Comparison of Figures 2 and 3 highlights that rhyme duration but not vowel duration correlates with syllable type (recall that tones 1 to 6 are unchecked and tones 7 to 10 are checked). Unchecked syllables are the relevant 'long duration' context for tones in Du'an Zhuang and checked syllables are the 'short duration' context. Syllable type correlated with rhyme duration and vowel length correlated with vowel duration, as expected. In unchecked syllables, rhyme duration and vowel duration are modulated independently of each other. However, in checked syllables, rhyme duration and vowel duration are equivalent due to the lack of sonorant codas. This results in a hierarchy of three different contexts for tones defined by rhyme duration: (CVV, CVVN, CVN) > CVVO > CVO, as can be seen in Figure 4. This result matches the situation seen in Thai (Zhang, 2001) and confirms the dominance of syllable type over vowel length (cf. Table 4).

One by-product of this three-tier hierarchy is that checked CVVO syllables have shorter rhyme durations than unchecked CVN syllables, but longer vowel durations. This would necessarily result in lengthening of the nasal coda in CVN syllables to allow a short vowel in a 'long duration' context. A CVVN syllable would then also need to have a relatively short nasal coda to accommodate the long vowel. These predictions are borne out, as shown in Figure 4.

In order to confirm the significance of this nasal coda duration effect, a linear mixed model was created for nasal coda duration with fixed effects for vowel length, tone, and their interaction, all of which were significant in a Type-III ANOVA (Tone:  $F(5,33.7) = 81.1; p < .001$ ; Vowel length:  $F(1,33.3) = 607.8; p < .001$ , Tone  $\times$  Vowel length:  $F(5,33.8) = 4.8; p = .002$ ). Vowel quality was not included as a random intercept as it did not provide a significant improvement to model fit ( $\chi^2 = 1.78; p = .180$ ). The coefficient of determination for the model was  $R^2 = .796$ . Post-hoc tests confirmed that nasal coda duration was longer

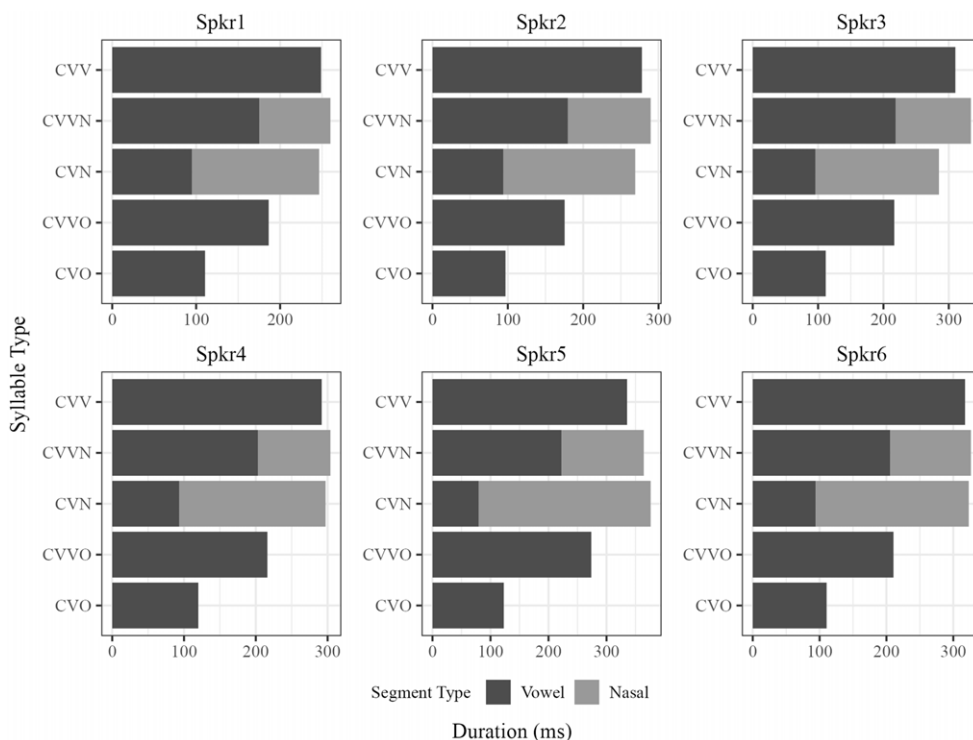


**Figure 3.** EMMs of vowel duration (ms) by tone with 95% confidence intervals.

following a short vowel (203 ms) than a long vowel (115 ms) ( $t = 24.6, p < .001$ ). Since there was a significant interaction between tone and vowel length, estimated marginal means for nasal coda duration are categorized by tone and vowel length in Figure 5.

Significance tests were performed for short vs. long vowels within each of the six unchecked tones, and all differences were significant ( $p < .001$ ). Nasal coda duration mirrored rhyme duration in that it was shorter in tones 1 and 6, but in addition, nasal duration was longer in tone 5 than all tones except tone 4 following a long vowel. However, there were no differences in nasal durations following a short vowel among tones 2, 3, 4 and 5, accounting for the interaction effect.

Finally, it can also be seen that vowel duration is longer in open CVV syllables than in in syllables closed by a coda, whether sonorant or obstruent. The fact that nasal or obstruent codas result in shortened vowel duration relative to open syllables is not surprising given the widely attested phenomenon of vowel shortening in the presence of coda consonants (Maddieson, 1985; Myers, 1987; Kubozono & Matsui, 2003; Maddieson, 2004). It also implies that rhyme duration is roughly equivalent in all unchecked syllables, regardless of the durations of the vowel or coda.



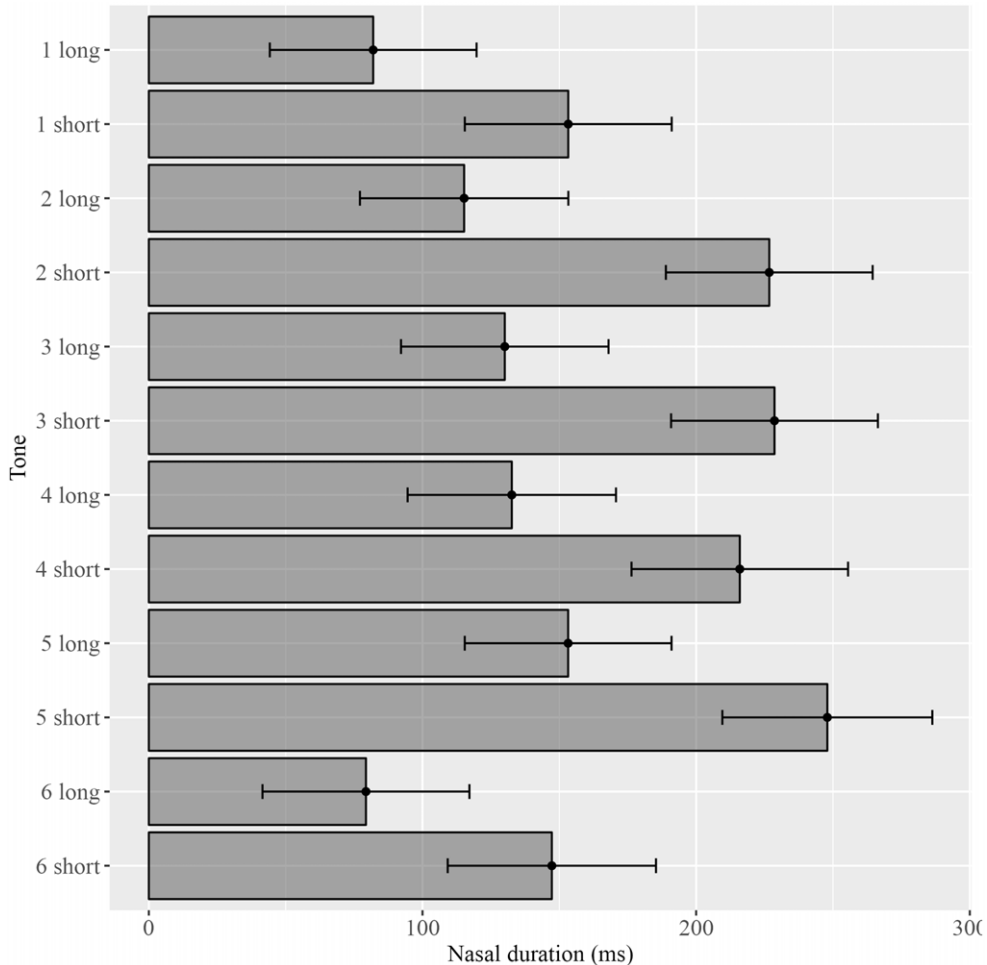
**Figure 4.** Rhyme Duration split into vowel and nasal duration (ms) by syllable type, one speaker per panel.

### 3.2 F0 Contours

Having established that rhyme duration correlates with the difference between syllable types, an analysis of  $f_0$ , the main phonetic correlate of tone is performed. Following this, interpretations regarding tonal complexity and its interaction with duration are made based on comparisons of the  $f_0$  contours across different syllable types.

To enable visualization of the  $f_0$  contours, plots of  $f_0$  in cents against normalized time are illustrated in Figure 6, one plot per speaker. The results show that five of the six tones in unchecked syllables are contour tones: Tones 1 and 6 are falling tones, tones 2 and 3 are rising-falling tones and tone 4 is a rising tone. Tone 5 is the only level tone and it is a mid tone. Among the checked syllables, the only clear contour tone is tone 10, which is a falling tone. The remaining three tones have larger amounts of variation between speakers and are more difficult to generalize across. However, if they are contour tones, all of them have reduced  $f_0$  excursions relative to their counterparts in unchecked syllables.

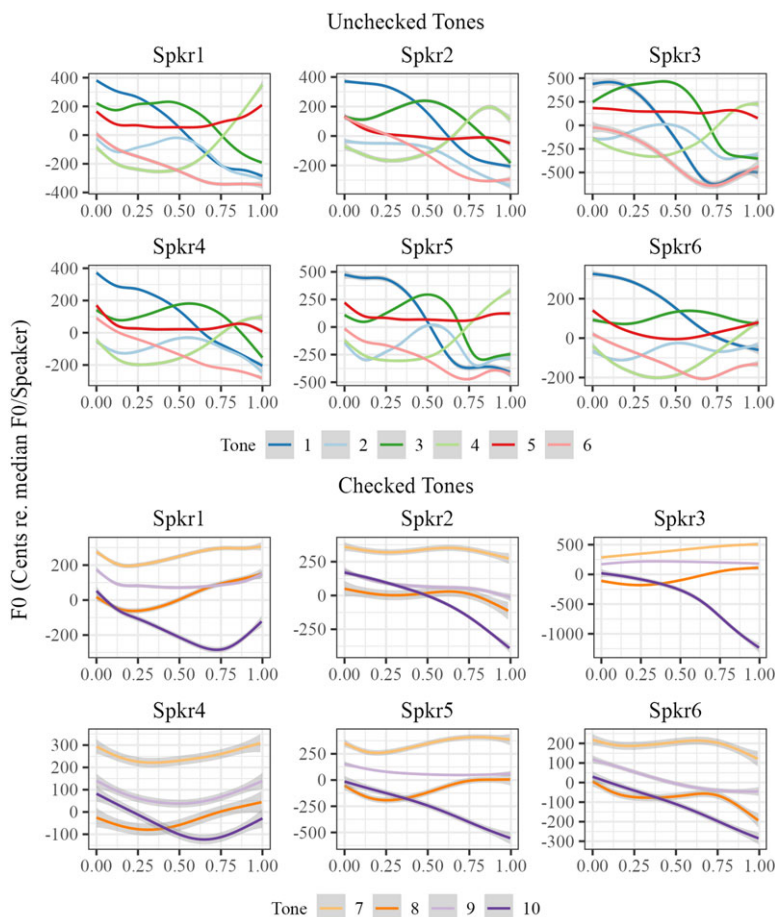
In order to generalize the  $f_0$  contour behavior across speakers, a GAMM model was constructed with both parametric and thin-plate regression smooth terms for tone. Random intercepts and random smooths for speaker, vowel quality, word and onset type using factor smooth interactions and an AR1 autoregressive error model were included. Coefficients of determination  $R_{uc}^2 = .792$ ,  $R_{cvo}^2 = .781$  and  $R_{cvo}^2 = .745$  were found respectively for the three models. Smooth splines, generated via `plot_smooth` of the `itsadug` package of R, are shown in Figure 7. There is a noticeable  $f_0$ -raising effect seen at the start of every tone. This effect is likely due to  $f_0$ -effects of onset consonants. There was an imbalance in the data set where the number of voiceless obstruents far outnumbered voiced consonant onsets (2,196 tokens



**Figure 5.** EMMs of nasal coda duration (ms) by tone and vowel length with 95% confidence intervals.

with voiceless obstruent onsets vs. 455 tokens with voiced onsets). These voiceless obstruents noticeably raised  $f_0$  for about the first 50 milliseconds of each syllable, whereas voiced segments had no local  $f_0$  perturbations. The random effect of onset type was removed from the plot in Figure 7 and this likely resulted in averaging across a biased data set, yielding the  $f_0$ -raising effect shown.

Next, the possibility of allotonic correspondences between tones in different syllable types is investigated. First, the six tones seen in unchecked syllables and the two tones in CVVO syllables (tones 9 and 10) are compared. Figure 7 shows that tone 9 is a mid-level tone, resembling tone 5, and tone 10 is a mid-falling tone, resembling tone 6. In order to check for allotonic correspondences, smooth scatterplots with the relevant tones were created via the ggplot function in R, like Figure 6. Two kinds of plots were created with differing measures of time on the x-axis: Ones with time normalization and ones without such normalization. Non-normalized time is useful in cases where allotones involve a truncated  $f_0$  contour and normalized time is useful in cases where  $f_0$  contours are preserved over a shorter duration, possibly at the expense of the  $f_0$  excursion extent.

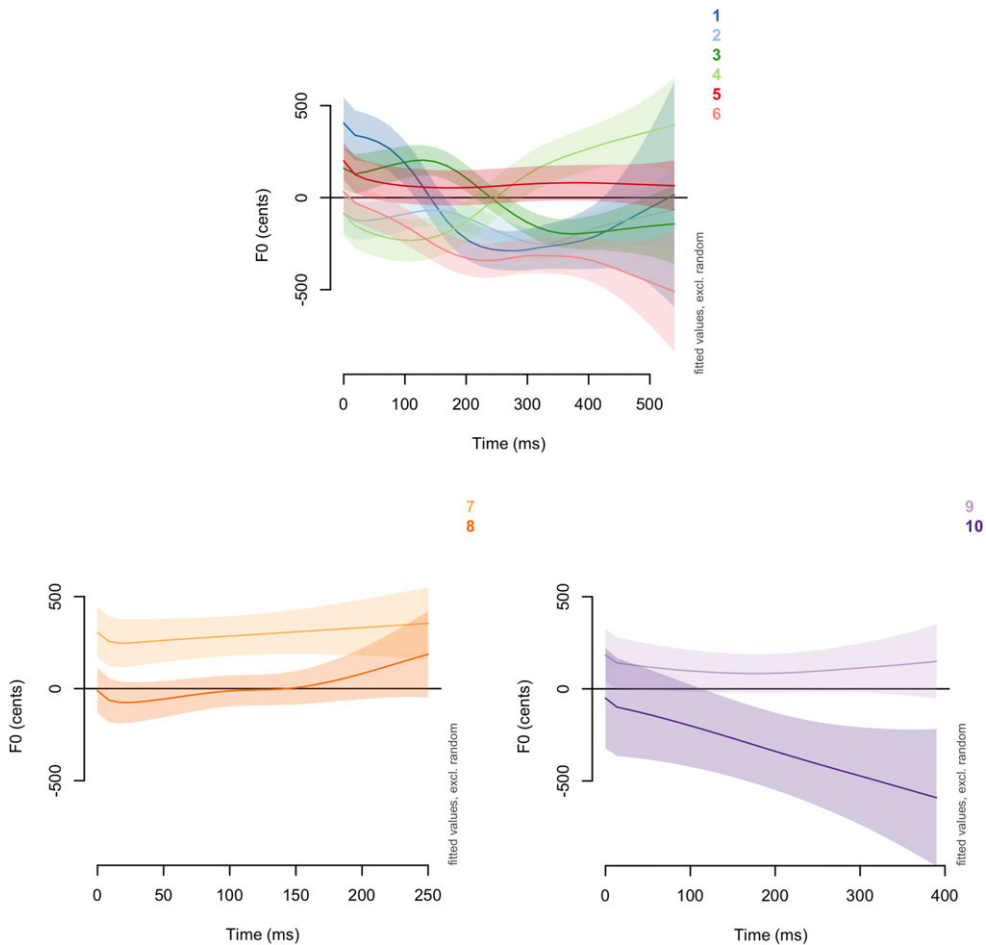


**Figure 6.** (Color online) Smooth spline scatterplots for  $f_0$  for each tone plotted against normalized time, one speaker per panel. Note that the  $f_0$  ranges vary per speaker.

Since absolute time yielded a closer match for tones 5 and 9, Figure 8 shows  $f_0$  against absolute time. Clear correspondences can be seen between mid-level tones 5 and 9, and also between the falling contour of tones 10 and 6. Similar plots using normalized time would not reveal the match between tone 10 and 6 as clearly, since the same  $f_0$  excursion is produced over a smaller proportion of the rhyme duration in tone 6 relative to tone 10.

Next, tones 7 and 8, which occur in CVO syllables are considered. Tone 7 is a high tone and may have a slight rising contour, while tone 8 is a mid tone, also with a possible rising contour. However, beyond these generalizations, the situation is less straightforward and variation between participants is considerable. As Figure 6 showed, speakers 2 and 6 have slight falling contours, while the other four speakers have slight rising contours for tone 8. Tone 7 may have a rising-falling contour similar to tones 2 and 3, but with a greatly reduced  $f_0$  excursion. An alternative possibility is that both tones 7 and 8 are level tones and that the slight contours seen in the smooth plots are artefacts of the smoothing.

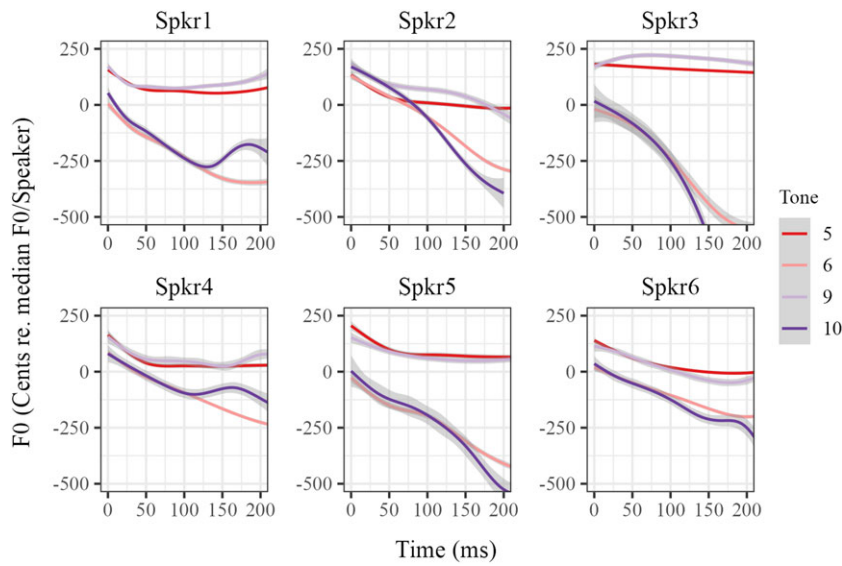
Considering tone 7 first, speakers 2, 5 and 6 have rising-falling  $f_0$  contours similar to those of tone 3, but with reduced  $f_0$  excursions. This can be seen in Figure 9, with normalized time. If this is a case of allotonic correspondence, it would qualify as tonal simplification in a shorter duration context (CVO) since the  $f_0$  excursions are greatly decreased



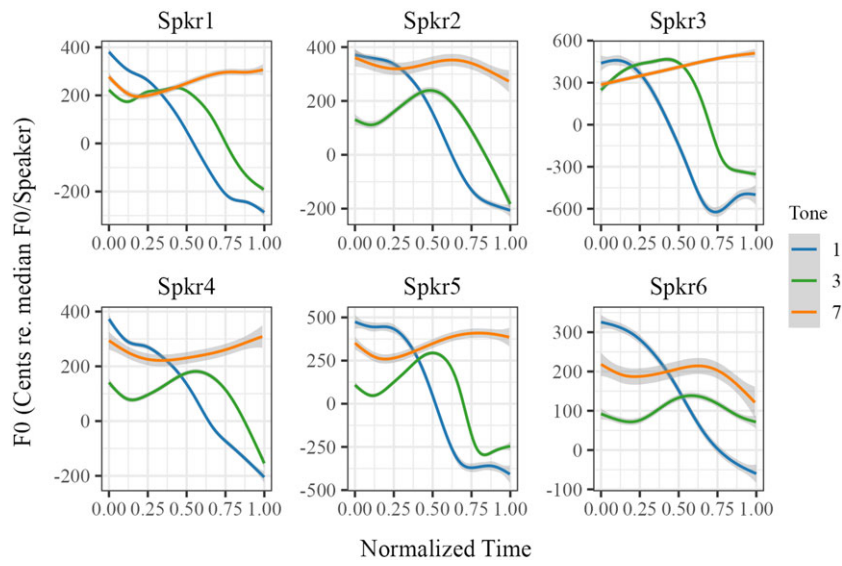
**Figure 7.** (Color online) F0 smooth splines for each tone in Du'an Zhuang plotted against time with 95% confidence intervals. Unchecked tones are on the top, CVO tones are on the bottom-left and CVVO tones are on the bottom-right. Significant differences between contours correspond to non-overlapping periods.

in tone 7. Alternatively, the tones may be high level tones. In either case, this involves a decrease in tonal complexity. In arguing for allotonic correspondences though, all three speakers also have significantly higher f0 for tone 7 relative to tone 3, which is not easily explained via allotonic correspondence.

Regarding speakers 1, 3 and 4, however, no such correspondence in contour shape with any of the unchecked tones is apparent in Figure 9. Instead, a slightly rising contour is produced for tone 7. However, inspection of plots showing f0 against non-normalized time in Figure 10 reveals that tone 7 matches the initial rise of tone 3 for these three speakers. The final fall that characterizes tone 3 is missing in tone 7. Thus, whereas speakers 2, 5 and 6 retained the contour shape, but reduced the f0 excursion, speakers 1, 3 and 4 retained the slightly rising f0 contour of the first part of the tone, simplifying the rising-falling contour to a rising contour. While speakers 1 and 3 show close correspondence in the f0 contour between tones 3 and 7, speaker 4 has an upward shift in f0 for tone 7 relative to tone 3, similar to what was seen for speakers 2, 4 and 6. This suggests that the f0 raising in tone 7 may be an independent adaptation that some speakers have.



**Figure 8.** (Color online) F0 contours for tones 5, 6, 9, and 10 over the initial 200 ms, one speaker per panel.

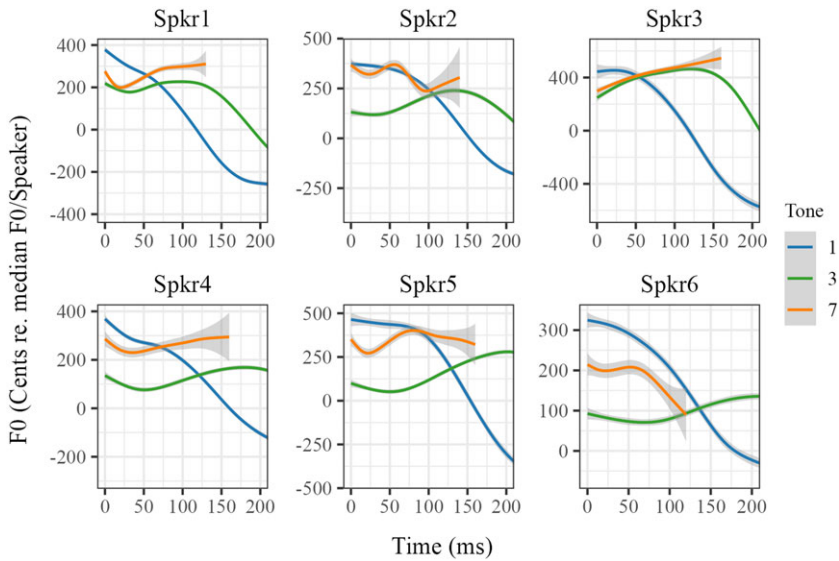


**Figure 9.** (Color online) F0 contours for tones 1, 3 and 7 by normalized time, one speaker per panel.

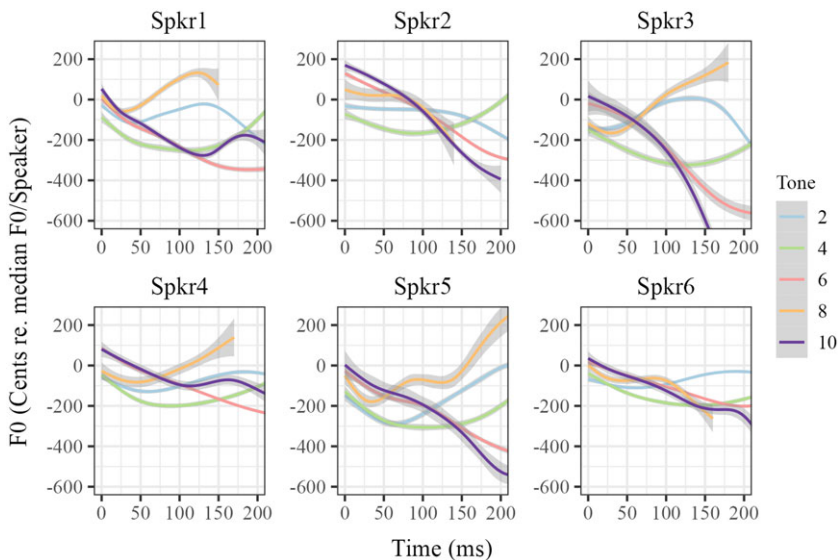
The evidence for allotonic correspondence between tones 7 and 3 is somewhat tenuous. However, regardless of this, tone 7 is either a contour tone (rising-falling for speakers 2, 5 and 6; simply rising for the others) with a much-reduced f0 excursion or it is a level high tone. In any case, both options are consistent with the prediction outlined in Section 1 for a tone in a shorter duration CVO context: In the face of the shorter duration context, speakers either reduce f0 excursions or produce less complex f0 contours.

Finally, tone 8 is considered. Like tone 7, the arguments for allotonic correspondences are somewhat unclear. However, some relevant observations can be made that point





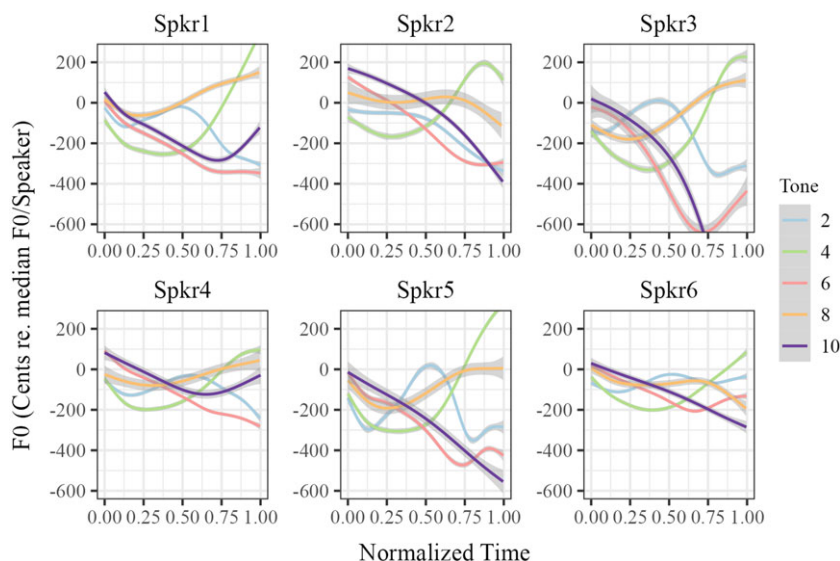
**Figure 10.** (Color online) F0 contours for tones 1, 3 and 7 by absolute time, one speaker per panel.



**Figure 11.** (Color online) F0 contours for tones 2, 4, 6, 8 and 10 by absolute time, one speaker per panel.

towards generalizations. Figure 11 shows f0 along absolute time for a selection of candidate allotones.

It can be observed that speakers 2 and 6 show a close correspondence between tones 6, 8 and 10, which are all mid-falling. The only difference is the slight rising-falling contour in tone 8 that is absent in tones 6 and 10; this rise-fall is similar to the rising-falling contour seen in tone 2. The remaining four speakers all produce tone 8 with a rising contour, similar to tone 4, but over a shorter duration. Figure 12 displays the same set of tones against



**Figure 12.** (Color online) F0 contours for tones 2, 4, 6, 8 and 10 by normalized time, one speaker per panel.

normalized time. For speakers 1, 3, 4 and 5, tone 8 is a rising tone with a smaller f0 excursion than tone 4.

Thus, there is some evidence of a correspondence between tone 8, 6, and 10 for two speakers and between tones 8 and 4 for four other speakers. In general, though, the evidence for allotonic correspondences for tones in the CVO context is weaker than the CVVO context. This may reflect the fact that the shortest duration context would naturally be the one where phonetic similarity is hardest to identify: The shorter the duration, the more the f0 excursions are expected to be altered, making allotonic correspondences less feasible. Regardless of whether these allotonic correspondences exist, tone 8 still exhibits reduced tonal complexity in a shorter duration (CVO) context, like tone 7: The rising f0 contour seen for four speakers has a reduced excursion amount relative to the longer duration (unchecked tone 4) context. Thus, the results show that Du'an Zhuang's tone system in shorter duration contexts, here CVO, is less tonally complex than the tones in the longer duration contexts (unchecked syllables).

#### 4. Discussion and conclusions

The first research question investigated whether rhyme duration correlated with syllable type. Recall that the tonal system is split by syllable type such that unchecked syllables allow six tones, but checked syllables allow four tones. The results showed that rhyme duration measurements correlated with this split. In particular, unchecked and checked syllables had distinct rhyme durations that corresponded exactly to the different tonal systems specific to each syllable type. Unchecked syllables had the longest rhyme duration and supported six tones; among checked syllables, there were four tones: CVVO syllables had an intermediate rhyme duration and supported two tones; CVO syllables had the shortest rhyme duration and supported the two other tones.

These results place Du'an Zhuang alongside Thai and Cantonese in that these languages have tonal systems that are sensitive to syllable type and not vowel length (Zhang, 2001, 2004). Recall that Table 4 summarized how syllable type (checked vs. unchecked) and vowel

**Table 6.** Hierarchy of rhyme duration difference results by combinations of syllable type and vowel length in Du'an Zhuang

Rhyme Duration		Syllable Type (Unchecked > Checked)	Vowel Length (Long > Short)
Longest duration	CVV, CVVN,	✓	✓
	CVN	✓	×
Medium duration	CVVO	×	✓
Shortest duration	CVO	×	×

Checkmarks correspond to cells predicted to have longer rhyme duration given their phonological status in each column.

length could interact regarding rhyme duration in a language. The rhyme duration difference between CVN and CVVO syllables was noted as being diagnostic: CVN syllables are unchecked but contain a short vowel; CVVO syllables are checked but contain a long vowel. Table 6 summarizes the results, showing that syllable type and not vowel length affects rhyme duration in Du'an Zhuang since CVN syllables had longer rhyme duration than CVVO syllables.

A separate investigation of vowel duration and nasal coda duration showed that vowel length contrasts exist independently, with unchecked syllables exhibiting coda durations that compensate for vowel duration differences. This allowed for equivalent rhyme durations among CVN and CVVN syllables. CVN syllables had shorter vowel duration but longer nasal codas, whereas CVVN syllables had longer vowels but shorter nasal codas. This situation is seen in other languages like Thai (Zhang, 2004; Morén & Zsiga, 2006). Among checked syllables, rhyme duration and vowel duration are equivalent since there are no sonorant codas. Therefore, syllable type can explain the three-way hierarchy seen in rhyme durations without any reference to vowel duration. What may appear on the surface to be an effect of vowel length on the tonal system in checked syllables can be explained via rhyme duration differences then. Therefore, while vowel length does differ between CVVO and CVO syllables, it is redundant and non-contrastive in these contexts.

After establishing this three-way hierarchy in rhyme duration, the tonal systems in each syllable type were investigated by measuring  $f_0$  and rhyme duration of each tone. This was done to address the second research question, which was to assess how tonal complexity relates to duration in Du'an Zhuang. There are five contour tones and one mid-level tone in unchecked syllables. This is reduced to one mid-level tone and one falling tone in CVVO syllables and to a pair of tones in CVO syllables that either have slight rising contours or consist of one level and one falling tone. Table 7 provides a summary, including a comparison with previous tonal descriptions. It adopts the subjective descriptive style of these previous descriptions, attempting to summarize the objective results described above.

Research question two asked whether tones in reduced duration contexts show signs of tonal complexity reduction. The first reduced duration context is in CVVO syllables. The two tones in the CVVO syllables, tone 9 and 10, had identical  $f_0$  contours to tones 5 (mid-level) and 6 (mid-falling). Based on this similarity, allotonic correspondences between these tone pairs are likely. The similarity of the  $f_0$  contours seen in Figure 8 suggests that the reduced duration in CVVO syllables was not accompanied by any phonetic simplifications (i.e. reduction in the  $f_0$  excursion).

However, it is notable that the two tones seen in CVVO syllables are allotones of the two least complex tones in the unchecked system, tones 5 and 6. Recall that tone 1 is a high-falling tone and tone 6 is a mid-falling tone. The total  $f_0$  excursion in tone 6 is less

**Table 7.** Description of the tonal system of Du'an Zhuang

	Tone	Description	Previous Description (when different)
Unchecked Tones	1	High falling	
	2	Low rising-falling	
	3	High rising-falling	
	4	Low rising	
	5	Mid level	
	6	Mid falling	Mid or low falling
Checked Tones	7	High rising or high level	High level
	8	Mid rising or mid falling	Mid level
	9	Mid level	Low rising
	10	Mid falling	Low falling

than that of tone 1, making it less complex than tone 1. Therefore, of the two falling tones, tone 6 is the least complex one and it is found in the reduced duration context (CVVO syllables). Likewise, tones 2 and 3 involve complex rising-falling contours and tone 4 is a rising tone, making it more complex than the two falling tones (Zhang, 2001, 2004). Therefore, the exclusion of these more complex tones (1, 2, 3 and 4) is expected in shorter duration contexts like CVVO syllables. Thus, the allotonic correspondences between tones 5 and 6 with 9 and 10 involve the two least tonally complex tones being preserved in the shorter duration CVVO context, as predicted.

However, the situation is less clear in the shortest duration context, CVO syllables (tones 7 and 8). As shown in Section 3, there was considerable variation between speakers on the f0 contours seen for tones 7 and 8. We concluded that tone 7 may be an unchecked allotone of the rising-falling tone 3 or it may alternatively be a level high tone. In the first case, all six speakers simplified the f0 contour in tone 7: Three speakers faithfully produced only the initial f0 rise, truncating the contour tone and excluding the final f0 fall; however, three other speakers faithfully produced a rising-falling contour, but with a reduced f0 excursion relative to tone 3. Therefore, the shorter duration context (CVO) saw speakers using two different strategies to reduce tonal complexity in tone 7 relative to tone 3. In the second case, where it is a high level tone, this would be straightforwardly predicted by Zhang’s theory since the level tone is the simplest possible tone. It would also match the situation in Thai and previous descriptions of Du’an Zhuang, where the two tones seen in CVO syllables are level tones.

Regarding tone 8, again variation was seen among speakers, with two speakers producing mid-falling tones, and the remaining four speakers producing rising tones. The mid-falling f0 contours produced by the first two speakers matched those seen in CVVO tone 10 and unchecked tone 6, suggesting an allotonic correspondence. The rising tone versions produced by the other four speakers had much-reduced f0-excursions compared to unchecked tone 4. Another possibility is that tone 8 may be an allotone of the rising-falling tone 2 for these speakers, analogous to the situation for tones 3 and 7. The speakers producing rising tones may be faithfully producing only the initial rise of tone 2 and truncating the final fall. If tones 7 and 8 are allotones of tones 3 and 2 though respectively, they both involve unexplained upward shifts of f0 relative to their unchecked allotones, and so this scenario is less likely. As can be seen from this discussion, it is difficult to definitively identify allotonic correspondences for tones in CVO syllables. Regardless of

this, both tones involve phonetic strategies where  $f_0$  contour excursions are reduced in the shortened duration context, as predicted.

While the above discusses the CVO and CVVO tonal systems relative to the unchecked system, it is informative to compare the checked syllable contexts to each other as well. This comparison yields a possible contradiction since CVO tones (tone 8 and possibly tone 7) may involve a more tonally complex rising contour compared to the relatively simple CVVO tones which are mid-level and mid-falling. This would be an exception to Zhang's (2001, 2004) generalization regarding tonal complexity and duration. That we see an exception may not be so surprising though, given that many diachronic processes happen despite being phonetically unnatural in a particular way. For example, the Proto-Tai long high vowel [i:] was shortened in modern Thai words with rising tone (Pittayaporn, 2009: 215), even though rising tone is a context that favors lengthening. However, if the CVO tones are actually high level tones, then no contradiction would exist, as a simpler level tone would be predicted in the shortest duration context.

Regarding the possibility that the CVO tones are both rising tones, we consider functional explanations: There is a good explanation for the lack of a vowel length contrast in checked syllables, but no such explanation for why CVO syllables but not CVVO syllables may host a rising tone. With rhyme durations already reduced among checked syllables, the contrastive load of duration would be very high if vowel length were contrastive. As we have seen above,  $f_0$  contours have smaller excursions in checked syllables and so it may be more difficult for speakers to distinguish these tones based only on  $f_0$ . By eliminating the vowel length contrast in checked syllables, this would essentially allow duration to act as a cue in addition to  $f_0$  for checked tones, thus boosting the contrast between, for example, tones 8 (short) and 10 (long) so that they are more easily recoverable for listeners. However, even so, this functional explanation does not explain why a given tone would prefer a short or long vowel, and so the fact that the more complex tones 7 and 8 only allow short vowels remains unexplained synchronically.

Instead, we must conclude that this is just the way the tone system evolved and that it represents a phonologically marked tone system. A historical picture emerges when considering Zhang et al.'s (1999: 24) correspondences between Proto-Tai (PT) tones and modern Zhuang tones: Proto-Tai tone \*C (tones 5 and 6, after the historical register split (Pittayaporn, 2009), which we abstract across here) are identical to PT tone \*DL (tones 9 and 10 here). Likewise, PT tone \*B (tones 3 and 4 here) are identical to PT tone \*DS (tones 7 and 8, here).

However, speakers did at least employ strategies to reduce tonal complexity. First, there is one exception to the historical correspondence described above: Recall that two of the six speakers exhibited a mid falling tone 8. If the historical picture above is correct, then these two speakers may have produced a cross-linguistically natural falling (like tone 10), rather than a rising tone, with the other four speakers exhibiting the standard rising tone for tone 8. This change from a more complex rise to a less complex fall would be a move in the direction expected by Zhang's theory of tonal complexity. Second, in tone 7, when speakers were faced with a complex rising-falling tone, they either reduced the extent of the  $f_0$  rise-fall or they simplified the complex contour to a rising contour. These phonetic simplifications can be seen as changes towards less complex tones, providing evidence that the situation is marked.

Related to this, while  $f_0$  contours were simplified in shorter duration contexts, it is notable that rhyme duration was relatively stable and was not altered as a strategy for tonal simplification in these shorter duration contexts in Du'an Zhuang. Recall the situation with Hausa, where falling tones had both a reduced  $f_0$ -excursion and also lengthened duration (Gordon, 1998; Zhang, 2001), both of which reduce complexity in CVO syllables. Du'an Zhuang, on the other hand, never employed duration changes as a strategy to reduce complexity. In fact, as mentioned above, the results for rhyme duration showed that tones with

greater complexity often had shorter, rather than longer, durations (although see Figure 3, where tone 8 did have a relatively longer inherent vowel duration than all other tones). These duration differences among checked syllables cannot be strategies for reducing tonal complexity then.

Instead, these duration differences seem to be inherent to the tones themselves. This possibility finds some support when considering the allotonic correspondence between tones 10 and 6: Recall that tone 6 had a relatively shorter duration than other unchecked tones, and so shortened duration may be inherent to the mid-falling tone, explaining why tone 10 has a relatively short duration itself. However, this allotonic explanation does not extend to the duration difference seen between tones 7 and 8 in CVO syllables. In this case, there was no plausible scenario that appealed to any of the unchecked allotone candidates of tones 7 and 8 that could explain duration differences, nor was there any functional explanation. This suggests that the tones in CVO syllables are not allotones of those in CVVO or unchecked syllables but are distinct tones with their own phonetic characteristics, including not only *f*<sub>0</sub> contour, but also perhaps rhyme duration in some cases. If this were true, it would add even more contrastive load to duration. However, the fact that rhyme duration is roughly equivalent between most tones within a syllable type does allow it some freedom to be used as a contrastive cue within each syllable type. In this way, rhyme duration defines three syllable types (unchecked, CVVO, and CVO) at a coarse-grained level, and within each syllable type, we speculate that it may be involved in finer-grained distinctions between tones. This speculation requires testing via perceptual tests to confirm in future research, however.

**Acknowledgments** We would like to thank the members of the Institute for Languages and Cultures of Asia and Africa joint research project entitled ‘Phonetic typology from cross-linguistic perspectives’ for their valuable feedback (phase 2, jrp000294). We would also like to thank the anonymous reviewers, whose suggestions have significantly improved this research. We are also indebted to Wentao Gu for input on an earlier version and Qiyuan Tang for his assistance in finding consultants. This work was funded by JSPS Kakenhi Grant-in-Aid for Young Scientists #15K16745 awarded to Jeremy Perkins and by JSPS Kakenhi Grant-in-Aid Category C #24K03872 awarded to Julián Villegas.

**Appendix A. Previous Descriptions of the Tonal System of Du’an Zhuang**

Note that Castro & Hansen only described the unchecked tones; examples are from Li (2011: 18).

	Tone	Zhang et al. 1999	Castro & Hansen 2010	Li 2011	Example Word
Unchecked Tones	1	42	High falling 53	42	厚 na:
	2	231	Low falling 21/231	232	黄 na:
	3	53	High rising-falling 453/343	55	脸 na:
	4	13	Low rising 214	213	舅舅 na:
	5	33	Mid level 33	33	骂 da:
	6	21	Mid falling 31	21	河 da:
Checked Tones	7	55	No description	55	肝 tap
	8	33	No description	33	墨 mak
	9	13	No description	213	搭配 tap
	10	21	No description	21	袜子 mat

## Appendix B. Word List

The word list in Table B1 shows each stimulus used in the experiment, presented in the order of the tone specifications in Qin (1996). IPA transcriptions from Qin (1996) are included. Note that for some prompts, more than one Zhuang word was listed in Qin (1996), and in those cases, both words were included from the single prompt. In cases where more than five tokens of a single IPA form by the same speaker resulted due to this, extra tokens were excluded (noted as ‘extra’ in the ‘# of Tokens Included’ column). When choosing which tokens to exclude on these grounds, sets of five repetitions of a word were kept preferentially, with incomplete sets excluded. In cases with more than one complete set or with no complete sets, then repetitions for the prompt with more tokens in the final data set were preferentially included. In cases where participants produced a word other than the expected Zhuang dictionary form, these were excluded from analysis, but were transcribed in the column ‘No Dictionary Match’; those produced with more than one syllable were not transcribed and are noted as ‘2- $\sigma$ ’.

**Table B1.** Word list

Word ID	Stimulus	Dictionary IPA (Qin 1996)	# of Tokens Included	# of Tokens Excluded			Skipped
				No Dictionary Match	f0-Related Error	Mandarin Chinese	
16	腿	ka:l	27	3 tinl	0	0	0
28	留 (as in 留给 he)	ɕe:l	27	2 hau3 l hen l	0	0	0
29	那 (as in 那个人)	te:l	22	5 lai2 l loi6 l pau4 l tanl	0	0	0
30	赊	θi:l	24	4 ɕi:5	2	0	0
22	租 (as in 租用)	ɕo:l	30	0	0	0	0
87	烧火 (as in 生火)	pau l	5	25 (2- $\sigma$ )	0	0	0
25	个 (as in 一个碗)	dan l	24	5 dak7	l	0	
2	啼 (as in 公鸡)	han l	15	15 heu6	0	0	0
1	真	ɕin l	30	0	0	0	0
5	脚	tin l	29	0	l	0	0
13	站	dum l	0	5 hum3 20 θoŋ2 5 θoŋ6	0	0	0
3	答	han l	30	0	0	0	0
10	煎 (as in 煎鱼)	ɕe:n l	28	0	2	0	0



**Table B1.** Continued

Word ID	Stimulus	Dictionary IPA (Qin 1996)	# of Tokens Included	# of Tokens Excluded			Skipped
				No Dictionary Match	f0-Related Error	Mandarin Chinese	
6	砖	ɕi:n1	30	0	0	0	0
11	月	du:n1	25	0	5	0	0
27	搬 (as in 东西)	pu:n1	28	0	2	0	0
40	茶	ɕa:2	29	0	1	0	0
41	鳃 (as in 骨鳃喉)	ka:2	14	1 man6	0	0	15
52	纺 (as in 纺线)	ta:2	5	6 ɕi:1 5 ta:3 5 ta:4 4 θa:1	0	0	5
45	黄牛	ɕu:n2	25	5 (2-σ)	0	0	0
176	淘 (as in 淘米)	ta:u2	5	20 θo:i5 5 θo:i6	0	0	0
46	黄蜂	tin2	22	8 to:5	0	0	0
37	句 (as in 一句话)	ɕon2	28	2 ɕe:6	0	0	0
145	下 (as in 下雨)	roŋ2 tok7	23 2 extras	5 tau3	0	0	0
50	柴	fun2	30	0	0	0	0
51	盆	pu:n2	30	0	0	0	0
4	跟	ruŋ2	0	29 ru:ŋ2 1 tem1	0	0	0
47	偏袒	ham2	0	5 ham4 9 hau3 5 he:n1 5 ho:1	1	0	5
33	房屋	ram2	30	0	0	0	0
35	钱	ɕi:n2	15	2 ŋan2 13 pi:n5	0	0	0
49	猜	tu:n2	5	20 ɕai1	5	0	0
76	等候	ɕa:3	30	0	0	0	0
144	打 (as in 打算盘)	ta:3	20 + 4 extras	1 ta:6 5 tak7	0	0	0
146	打 (as in 打球)	ta:3 tuk7	5 + 15 extras 3	2 tuk8	5	0	0



Table B1. Continued

Word ID	Stimulus	Dictionary IPA (Qin 1996)	# of Tokens Included	# of Tokens Excluded			Skipped
				No Dictionary Match	f0-Related Error	Mandarin Chinese	
64	从 (as in 从南宁来)	ta:3 ɕoŋ2	5 + 10 extras 4	5 tau3 5 ta:5	1	0	0
77	乳房	ɕi:3	30	0	0	0	0
66	值 (as in 值多少钱)	ti:3	30	0	0	0	0
67	躲	do:3	30	0	0	0	0
164	凸	mo:3	9	1 mu:n4 5 tu:5			15
73	米粉	fan3	29	0	1	0	0
56	捏 (as in 捏鼻子)	ɲan3	15	5 bu:n7 5 ɕau4	0	0	5
54	穿 (as in 穿衣)	tan3	30	0	0	0	0
63	上 (as in 上山)	hən3	30	0	0	0	0
74	本 (as in 一本书)	pən3	29	0	1	0	0
59	短	tin3	30	0	0	0	0
53	吞	tu:n3	30	0	0	0	0
115	建 (as in 建房)	hu:n ku:6	7 10	1 kje:n4 12 ku:4	0	0	0
61	田	he:n3	30	0	0	0	0
75	扁	pe:n3 be:l	19 5	5 pe:n2	1	0	0
60	浅 (as in 浅水)	ɕi:n3	15	10 fe:u6	5	0	0
83	*疏	ɕa:4	5	10 kjai1	0	0	15
85	知道	ro:4	30	0	0	0	0
91	剁 (as in 剁肉)	to:4	29	0	1	0	0
108	最 (as in 最高)	θe:i4	30	0	0	0	0
88	尘土	fan4	24	5 nam6	1	0	0

**Table B1.** Continued

Word ID	Stimulus	Dictionary IPA (Qin 1996)	# of Tokens Included	# of Tokens Excluded			Skipped
				No Dictionary Match	f0-Related Error	Mandarin Chinese	
80	溢	ron4	0	5 ji:4 5 mu:n4 5 mu:t10 10 rim1	0	0	5
89	啃 (as in 骨头)	hen4	29	1 (2-σ)	0	0	0
90	陪伴	pun4	30	0	0	0	0
107	借	ci:5	30	0	0	0	0
128	淋	co:5	0	4 lin6 5 ran2 4 ru:t10 5 θam4 5 θan1 5 θo:4	2	0	0
111	马蜂	to:5	30	0	0	0	0
70	吓唬	hu:5	0	5 ja:5 10 lo:4	5	0	10
109	兔子	tu:5	10	20 tu:4	0	0	0
113	够	kau5	30	0	0	0	0
112	旋转	pan5	14	16 ci:n5	0	0	0
24	盖 (as in 盖被子)	hum5	21	5 kai4 4 kwop8	0	0	0
96	鹅	ha:n5	29	1 pit7	0	0	0
92	炭	ta:n5	25	0	5	0	0
97	寸	co:m5	20	5 cik7 5 θok7	0	0	0
99	切 (as in 切菜)	ro:n5	22	5 dat7 3 kat7	0	0	0
98	顿 (as in 骂一顿)	to:n5	20	3 cam1 4 cam5	2	1	0
119	河	ta:6	29	0	1	0	0
124	种子	ce:6	30	0	0	0	0
125	核	ŋwe:6	4	7 ce:6 5 hu:6	1	0	13
9	路	lo:6	29	0	0	0	1

Table B1. Continued

Word ID	Stimulus	Dictionary IPA (Qin 1996)	# of Tokens Included	# of Tokens Excluded			Skipped
				No Dictionary Match	f0-Related Error	Mandarin Chinese	
123	漏 (as in 漏雨)	ro:6 lau6	15 10	4 lau4	0	0	1
127	豆	tu:6	30	0	0	0	0
169	是	θe:i6	28	0	2	0	0
114	修理	ɕu:i6	24	6 (2-σ)	0	0	0
118	赞扬	han6	19	10 (2-σ)	1	0	0
132	粪	pun6	25	3 hai4 1 ku:n3 1 tan5	0	0	0
116	赚	ɕam6	14	10 to:2 1 θu:n5	5	0	0
130	万	fam6	29	0	1	0	0
117	汗	ham6	29	0	1	0	0
131	半	to:n6	25	5 pu:n5	0	0	0
149	滴 (as in 一滴水)	ɕak7	22	4 nik7	4	0	0
129	追	kjap7 ɕan6	25 5	0	0	0	0
147	摘 (as in 摘花)	bat7	20	5 bat8	0	0	5
137	七	ɕat7	30	0	0	0	0
136	紧 (as in 绷紧)	dat7	24	1 kan3	5	0	0
20	一些	nat7	2	15 (2-σ) 8 ɕak7 4 pu:l	1	0	0
148	扫 (as in 扫地)	pat7	30	0	0	0	0
138	菌子	rat7	0	5 (2-σ) 5 kjin4	0	0	20
135	淡 (as in 不咸)	ɕit7	29	0	1	0	0
152	鸭子	pit7	30	0	0	0	0
151	弹 (as in 用指)	tit7	1	5 mi:m6 5 te:t9 14 tik7	0	0	5

**Table B1.** Continued

Word ID	Stimulus	Dictionary IPA (Qin 1996)	# of Tokens Included	# of Tokens Excluded			Skipped
				No Dictionary Match	f0-Related Error	Mandarin Chinese	
155	落 (as in 树叶落)	tok7	30	0	0	0	0
133	打 (as in 打铁)	tuk7 ta:3	4 16 extras	0	10	0	0
143	捕 (as in 鱼)	tuk7	2	4 cam3 15 kap8 4 lau5	0	0	5
72	朵 (as in 一朵花)	tup7	13	16 tu:3	0	0	1
141	蕨草	kut7	0	5 rum3	0	0	25
159	贼	ɕak8	30	0	0	0	0
157	打 (as in 用帚)	map8	0	11 pat7 1 ta:2 8 ta:3 4 ta:5	6	0	0
171	神	pat8	10	1 ɕin2 5 fu:n3 5 θi:n1 5 θun6	0	0	4
158	剪	rat8	12	1 ɕa:4 3 kat7 9 tat7	5	0	0
57	急 (as in 水流急)	kip8	0	14 kan3 14 rau2	2	0	0
172	蝉	pit8	17	5 pit7 1 tin2	1	0	6
38	被子	mok8	25	4 ten2	0	0	1
165	读	tok8	30	0	0	0	0
170	稠 (as in 稠粥)	ku:k8	0	14 hau5 7 ni:u1 5 noŋ2	1	0	3
168	熟 (as in 熟肉)	ɕuk8	24	0	5	0	1
156	打 (as in 用拳头)	tup8	10	5 mi:n6 15 ta:3	0	0	0
178	晒	ta:k9	25	5 ɕai4	0	0	0
175	刷子	ɕa:t9	15 + 4 extras	1 ɕa:2 4 ɕa:6	1		5
182	擦 (as in 擦桌子)	ɕa:t9 wu:t9	10 + 5 extras 5	7 mu:t10 3 pat7	0	0	0

Table B1. Continued

Word ID	Stimulus	Dictionary IPA (Qin 1996)	# of Tokens Included	# of Tokens Excluded			Skipped
				No Dictionary Match	f0-Related Error	Mandarin Chinese	
174	蜚 (as in 蜂蛰人)	da:t9	30	0	0	0	0
181	削 (as in 削尖)	ta:t9	0	29 tat7	1	0	0
26	隔开	ke:k9 he:n1	11 16	3 hai1	0	0	0
103	骂	we:t9	30	0	0	0	0
183	节日	ɕit9	20	5 (2-σ)	0	0	5
184	啄 (as in 啄树)	to:t9	29	1 fok7	0	0	0
185	揭开	ku:t9	8	5 ɕit10 1 hai1 5 hai2 4 hai3 1 hen1 1 kit10 4 piŋ3	0	0	1
180	脱 (as in 脱衣)	tut9	30	0	0	0	0
188	绳子	ɕak10	28	0	0	0	2
142	烧	la:k10	0	24 pau1 5 ta:t9	1	0	0
189	量 (as in 比量)	ta:k10	11	5 kjau4	0	5	5
		rau1	4				
186	打 (as in 打谷)	fat10	0	1 map8 5 mi:n6 18 ta:3 1 tup8	5	0	0
191	女阴	ɕet10	25	0	0	0	5
192	吝啬	ket10	25	5 (2-σ)	0	0	0
196	地方	tuk10	0	30 (2-σ)	0	0	0
195	扛	ku:t10	22	4 ru:m3 1 ta:5 1 ?am5	2	0	0
197	臭虫	ɣut10	0	11 (2-σ) 5 hau1	1		13
193	翅膀	fu:t10	30	0	0	0	0
<b>Total</b>	<b>133 items</b>		<b>2647 + 56 extras</b>	<b>965</b>	<b>114</b>	<b>6</b>	<b>202</b>

## References

- Abramson, Arthur. 1962. The vowels and tones of standard Thai: Acoustical measurements and experiments. *International Journal of American Linguistics* 28(2), part II.
- Baayen, R. Harald, Jacolien van Rij, Cecile de Cat & Simon Wood. 2018. Autocorrelated errors in experimental data in the language sciences: Some solutions offered by generalized additive mixed models. In D. Speelman, K. Heylen, D. Geeraerts (eds.), *Mixed-effects regression models in linguistics. Quantitative methods in the humanities and social sciences*. Cham: Springer. [https://doi.org/10.1007/978-3-319-69830-4\\_4](https://doi.org/10.1007/978-3-319-69830-4_4)
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48.
- Boersma, Paul & David Weenink. 2018. Praat: Doing phonetics by computer (version 6.2.14).
- Brunelle, Marc & Joshua Finkeldey. 2011. Tone perception in Sgaw Karen. *17th International Congress of Phonetic Sciences (ICPhS XVII)*, Hong Kong, 372–375.
- Castro, Andy & Bruce Hansen. 2010. Hongshui He Zhuang dialect intelligibility survey. *SIL Electronic Survey Reports* 2010–025.
- Coupe, Alexander R. 2014. Strategies for analyzing tone languages. *Language Documentation & Conservation* 8, 462–489.
- DiCanio, Cristian T. 2009. The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association* 39(2), 162–188.
- Duanmu, San. 1990. A formal study of syllable, tone, stress and domain in Chinese languages. Ph.D dissertation, MIT.
- Esposito, Christina M. 2012. An acoustic and electroglottographic study of White Hmong tone and phonation. *Journal of Phonetics* 40(3), 466–476.
- Fu, Qian-Jie & Fan-Gang Zeng. 2000. Identification of temporal envelope cues in Chinese tone recognition. *Asia Pacific Journal of Speech, Language and Hearing* 5(1), 45–57.
- Gandour, Jack. 1977. On the interaction between tone and vowel length: Evidence from Thai dialects. *Phonetica* 34(1), 54–65.
- Garellek, Marc, Patricia Keating, Christina M. Esposito & Jody Kreiman. 2013. Voice quality and tone identification in White Hmong. *The Journal of the Acoustical Society of America* 133(2), 1078–1089.
- Gordon, Matthew. 1998. The process-specific nature of weight: The case of contour tone restrictions. In Kimary N. Shahin, Susan Blake, and Eun-Sook Kim (eds.), *Proceedings of WCCFL 17*. GLSA Publications, Amherst, MA.
- Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29(4), 383–406.
- GZZW (Guangxi Zuangzu Zizhiqun Shaoshuminzu Yuyanwenzizongzuo Weiuyuanhui & Zhuanghanyingcidian Weiuyuanhui). 2005. *壮汉英词典 [Guengh Gun Yingh swzdenj] (Zhuang-Chinese-English dictionary)*. Beijing: Minzu chubanshe.
- Kawahara, Hideki. 2006. STRAIGHT, exploitation of the other aspect of VOCODER: Perceptually isomorphic decomposition of speech sounds. *Acoustical Science and Technology* 27(6), 349–353.
- Kawahara, Hideki, Ikuyo Masuda-Katsuse & Alain De Cheveigné. 1999. Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based f0 extraction: Possible role of a repetitive structure in sounds. *Speech Communication* 27(3), 187–207.
- Kuang, Jianjing. 2017. Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice. *The Journal of the Acoustical Society of America* 142(3), 1693–1706.
- Kubozono, Haruo and Michinao Matsui. 2003. Phonetic vs. phonological control of speech: Closed syllable vowel shortening in Japanese dialects. *Proceedings of ICPhS XV*, 2429–2432.
- Kullavanijaya, Pranee. 2001. A study of lexical variation in seven Zhuang dialects. *Essays in Tai Linguistics*, 229–258.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. Package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13), 1–26.
- Lehiste, Ilse. 1976. Influence of fundamental frequency pattern on the perception of duration. *Journal of Phonetics* 4(2), 113–117.
- Lenth, Russell V. 2023. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.6. <https://CRAN.R-project.org/package=emmeans>
- Li, Xu-Lian. 2011. *都安壮语形态变化研究 [Du'an Zhuang yu xingtai bianhua yanjiu] (Du'an Zhuang morphological change research)*. Beijing: Minzu chubanshe.
- Liu, Siyun & Arthur G. Samuel. 2004. Perception of Mandarin lexical tones when f0 information is neutralized. *Language & Speech* 47(2), 109–138.
- Maddieson, Ian. 1985. Phonetic cues to syllabification. Edited by Victoria A. Fromkin. *Phonetic linguistics: Essays in honor of Peter Ladefoged*, 203–221. Orlando, FL: Academic Press.
- Maddieson, Ian. 2004. Timing and alignment: A case study of Lai. *Language and Linguistics* 5(4), 729–755.

- Morén, Bruce & Elizabeth Zsiga. 2006. The lexical and post-lexical phonology of Thai tones. *Natural Language and Linguistic Theory* 24(1), 113–178.
- Myers, Scott. 1987. Vowel shortening in English. *Natural Language & Linguistic Theory* 4(5), 485–518.
- National Bureau of Statistics. 2012. 中国 2010 年人口普查分县资料 [Zhongguo 2010 nian renkoupu chafen xian ziliao] (Tabulation on the 2010 population census of the People's Republic of China by county). Zhongguo tongji chubanshe.
- Ohala, John J. & William G. Ewan. 1973. Speech of pitch change. *The Journal of the Acoustical Society of America* 53, 345.
- Perkins, Jeremy, Seunghun J. Lee & Julián Villegas. 2016. An interplay between f0 and phonation in Du'an Zhuang tone. *Proceedings of the 5th International Symposium on Tonal Aspects of Languages*, 56–59.
- Pittayaporn, Pittayawat. 2009. The Phonology of Proto-Tai. Ph.D. dissertation, Cornell University.
- Qin, Guo-Sheng. 1996. 壮语方言概论 [Zhuang yu fangyan gailun] (An introduction to dialects of Zhuang language). Nanning: Guangxi minzu chubanshe.
- Qin, Xiao-Hang. 1995. 现代壮语 [Xian dai Zhuang yu] (Modern Zhuang Language). Beijing: Minzu chubanshe.
- R Core Team. 2019. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Shen, X. 2010. The acoustic performances of glottal stop. *Studies in Language and Linguistics* 30, 35–39.
- Shue, Yen-Liang, Patricia Keating, Chad Vicens & Kristine Yu. 2011. VoiceSauce: A program for voice analysis. 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong, 1846–1849.
- Sósokuthy, Márton. 2017. Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction. arXiv: 1703.05339.
- Tsanas, Athanasios, Matías Zañartu, Max A. Little, Cynthia Fox, Lorraine O. Ramig & Gari D. Clifford. 2014. Robust fundamental frequency estimation in sustained vowels: Detailed algorithmic comparisons and information fusion with adaptive Kalman filtering. *The Journal of the Acoustical Society of America* 135(5), 2885–2901.
- van Rij, Jacolien, Martijn Wieling, R. Harald Baayen & H. van Rij. 2022. Itsadug: Interpreting Time Series and Autocorrelated Data using GAMMs. R package version 2.4.1.
- Villegas, Julián, Seunghun J. Lee, Jeremy Perkins & Konstantin Markov. 2020. Prediction of creaky speech by recurrent neural networks using psychoacoustic roughness. *IEEE Journal of Selected Topics in Signal Processing* 14(2), 355–366.
- Wei, Qing-Wen & Guo-Sheng Qin, 1980. 壮语简志 [Zhuang yu jianzhi] (A Brief Description of the Zhuang Language). Beijing: Minzu chubanshe.
- Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.
- Wood, Simon. 2017. *Generalized Additive Models: An introduction with R*, 2nd ed. New York: Chapman and Hall/CRC Press.
- Yip, Moira. 1982. Against a segmental analysis of Zahao and Thai: A laryngeal tier proposal. *Linguistic Analysis* 9(1), 79–94.
- Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.
- Yu, Alan C. L. 2003. Contour tone induced lengthening in Cantonese. Edited by M. J. Sole, D. Recasens and J. Romero. *The 15th International Congress of Phonetic Sciences*, 2381–2384. Barcelona, Spain.
- Yu, Alan C. L. 2010. Tonal effects on perceived vowel duration. *Lab Phonology* 10(4), 151–168.
- Zhang, Jun-Ru, Min Liang, Jue-Ya Ouyang, Yi-Qing Zheng, Xu-Lian Li & Jian-Xian Xie 1999. 壮语方言研究 [Zhuang yu fangyan yanjiu] (A Study of Zhuang dialects). Chengdu: Sichuan minzu chubanshe.
- Zhang, Yuan-Sheng & Xiao-Hang Qin. 1993. 现代壮汉语比较语法 [Xiandai zhuang hanyu bijiao yufa] (A Modern Comparative Grammar of Zhuang and Chinese). Beijing: Zhongyang minzu xueyuan chubanshe.
- Zhang, Jie. 2001. The effects of duration and sonority on contour tone distribution: Typological survey and formal analysis. Ph.D. dissertation, UCLA.
- Zhang, Jie. 2004. The role of contrast-specific and language-specific phonetics in contour tone distribution. Edited by B. Hayes, R. Kirchner and D. Steriade. *Phonetically based phonology*, 157–190. Cambridge, UK: Cambridge University Press.