Tone sandhi, prosodic phrasing, and focus marking in Wenzhou Chinese

Scholz, F.

Chapter 2

Wenzhou Chinese phonetics and phonology sketch

2.1 General introduction

2.1.1 Area and intelligibility of the Wenzhou dialect

Wenzhou Chinese is a dialect of Chinese which is spoken in the Wenzhou urban area, home to approximately 9.12 million people.[1] Of the varieties spoken in the area, the one spoken in the central district of Wenzhou (Lucheng) is commonly regarded as most representative. It is more or less mutually intelligible with the other varieties spoken in and around the city, although speakers from Lucheng can usually point out the idiosyncrasies of speakers from the outskirt districts such as Longwan and Ouhai.

The Wenzhou dialect, as the dialect of the Lucheng area in Wenzhou will be referred to hereafter, is considered to be representative of the Oujiang dialect subgroup (Fu & Fang 1985), which belongs to the larger dialect group of Wu dialects. The term "dialect" in this context is commonly used by Chinese linguists to describe the different sub-groupings of Sinitic languages within China. In Western linguistic terms, the subgroups would commonly be classified as individual languages, partially because mutual intelligibility between the "dialects" can be low (Tang & van Heuven 2007, 2009). This thesis will follow the nomenclature of "dialect" when talking about the subgroups of Chinese, and refer to the Chinese standard language as Standard Chinese.

2.1.2 Previous descriptions of the Wenzhou dialect

The Wenzhou dialect of Chinese has been included in recent cross-dialectal descriptions in Chinese, such as the description by Hou (1998) in the *现代汉语方言音库 [Modern Chinese Dialect Sound Database]*-Series, or the *当代吴语研究 [Studies in the Contemporary Wu Dialects]*-Volume by Qian (1992). These accounts describe the dialect from a comparative perspective, with Standard Chinese as a reference point. Recent descriptions which focus

[1] Data from the 2010 6th national population census of China, available at http://www.wzstats.gov.cn/2010rkpc/infoshow.asp?id=4336.

specifically on the Wenzhou dialect have been published by You (2002) and Zhengzhang (2008).

Modern Western linguistic accounts of Wenzhou Chinese are often concerned with subparts of the phonological system, such as the account of the tone sandhi phonology in Chen 2000, or the phonetic descriptions of the tone and tone sandhi patterns in Rose 2000, 2001, 2002, 2004. Both of these sources derive their observations from recordings made in the 1980s with middle-aged speakers, and are therefore likely to represent an older state of the dialect.

The current account of the Wenzhou dialect is based on the speech of young speakers from Lucheng Wenzhou. All speakers recorded for this phonology sketch, as well as the thesis at large, were between 18 and 29 years of age.[2] Due to ongoing language contact with Standard Chinese, which the young speakers learned from an early age, their way to speak Wenzhou differs quite substantially from the above-mentioned descriptions in many respects. For this reason, the upcoming chapter attempts to give an overview of the young speaker's segmental and tonal inventory, so as to lay a basis for the descriptions in the remainder of the thesis.

In the following, the sound inventory of Wenzhou will be illustrated with example words given in Chinese characters (Hanzi), as well as in broad IPA transcriptions. For the young speakers, the immediate association with Chinese characters is Standard Chinese, and when asked to pronounce an individual character in Wenzhou dialect, they are sometimes unsure of the "correct" pronunciation. In this description, it was attempted to select illustrative characters which would be recognized by most speakers to be associated with the given pronunciation. Most of the examples in the current sketch were picked from the words that were used as illustrations in previous published descriptions of Wenzhou.

Throughout this thesis, the lexical tones of Wenzhou will be represented by Chao numbers, following the numerical description of Wenzhou in You 2002. The tone trajectories and characteristics themselves will be discussed in section 2.4.

[2] More detailed information on the speakers can be found in the individual chapters. The current sketch is mainly based on recordings with two female speakers, who were both 25 years old. All illustrations of F0 and spectral values in this sketch were produced from recordings with either of the two speakers.

2.2 Phoneme inventory

2.2.1 Consonants

Table 2.1: Overview of consonantal phonemes and their place and manner of articulation in Wenzhou Chinese.

	Bilabial	Labiodental	Alveolar	Alveolo-palatal	Velar	Glottal
Plosive	p^h p b		t^h t d		k^h k g	(ʔ)
Affricate			ts^h ts dz	$tɕ^h$ tɕ dʑ		
Nasal	m		n	ɲ	ŋ	
Fricative		f v	s z	ɕ		h ɦ
Approximant		ʋ		j		
Lateral approximant			l			

2.2.1.1 Plosives

(1)	Citation forms	Hanzi	Translation
a.	*p^ha42*	派	‘to send, dispatch’
b.	*pa42*	扮	‘to dress up’
c.	*ba31*	排	‘to arrange’
d.	*t^ha42*	太	‘excessively, too’
e.	*ta42*	带	‘belt, zone’
f.	*da31*	谈	‘to talk, chat’
g.	*k^ha42*	快	‘fast’
h.	*ka42*	界	‘boundary’
i.	*gau24*	厚	‘thick’
j.	*ʔa35*	矮	‘short, low’

The plosive onsets of Wenzhou differentiate four different places of articulation: bilabial [p^h, p, b], alveolar [t^h, t, d], velar [k^h, k, g], and glottal [ʔ]. In terms of voicing, Wenzhou displays a three-way voicing distinction in plosives and affricates. This voicing distinction is correlated with tonal register in a way that the high register tones (H) only occur on syllables with voiceless and voiceless aspirated onsets, and the low register tones (L) only occur on voiced onset syllables. The minimal pair examples in (1) are limited by this tonal differentiation.

The exact phonetic nature of the three-member voicing contrast in Wenzhou and many other Wu dialects has attracted research interest for a long time. Instrumental studies going back to Chao 1928 have shown that none of the three "voicing" states is actually characterized by vocal fold vibration in the onset consonant when it is in word-initial position. Furthermore, it has been established for related dialects that in initial position, the difference of VOT between the so-called "voiced" and voiceless unaspirated onsets is negligible (Cao 1982; Shi 1983).

In the aerodynamic study by Cao & Maddieson (1992), it was shown that while there is no phonetic "voicing" contrast in Wenzhou onsets in initial position, the onsets that are traditionally described as "voiced" and "voiceless" do nonetheless differ in tenseness and aperture of the vocal folds. This causes a phonation difference at the release of the stops, and some breathiness at the onset of the following vowel, which aligns with the traditional terminology of "murmured" vowels (Duanmu 2007). According to Rose (2001), this difference conditions the so-called "depression" effect, which lowers the F_0 onset of the tonal contour following the "voiced" onset. Crucially, he argues that in Wenzhou this depression effect is independent of the register of the tone, and can also be found on tonal contours which have been changed to e.g. high falling in a tone sandhi environment.

Therefore, when talking about "voicing" in Wenzhou, several connected phenomena need to be kept apart. Phonetic voicing, as defined by a difference in VOT, is only found in word-medial position between the "voiced" and "voiceless" onset consonants and affricates, but perceptually neutralized in word-initial position. At the same time, the "depression" effect of the "voiced" onsets onto the beginning of the F_0 contour is present in both contexts. Finally, the limitation of syllables with "voiced" onsets to only carry low register tones is sometimes eliminated in tone sandhi context, but independent of the "depression

effect". For further discussion of the tone system of Wenzhou, the reader is referred to section 2.4 of this thesis.

In the examples in (1), the final example has been transcribed with a glottal stop [ʔ] onset. In Chinese linguistics, the existence of glottal stops in Chinese dialects is subject to debate for reasons of diachronic reconstruction (Zhang 2006). For this reason, syllables like those in (1j) have been alternatively transcribed as onsetless [a] (Zhengzhang 2008) or as having a zero onset [0a] (You 2002; You & Yang 1998) in Wenzhou, while other descriptions used the glottal stop onset [ʔa] (Qian 1992; Rose 2001).

For the young speakers described in this sketch, it could be observed that their realization of examples like (1j) can vary between realizations with a clear glottal contraction at the beginning, or with a more breathy transition into the vowel without glottal constriction (see Figure 2.1). This inter-speaker variation could be a reason for the different transcriptions of this sound. Since there is no phonemic alternation between onsetless syllables and those with a glottal stop onset, the glottal stop is set in parentheses in Table 2.1. For this sketch, in spite of the inter-speaker variation, the glottal stop will be consistently used to transcribe otherwise onsetless syllables.

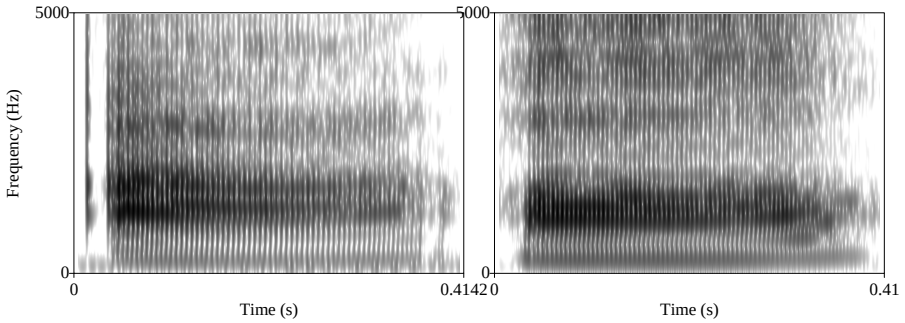


Figure 2.1: Realizations of example 1j (ʔa35) with (left) and without (right) glottal stop onset by two female speakers.

2.2.1.2 Affricates

(2)	Citation forms	Hanzi	Translation
a.	*ts^{h}a33*	差	'differ, wrong'
b.	*tsa42*	蘸	'to dip in liquid'

c.	*dza31*	惭	‘ashamed’
d.	*tɕʰa313*	雀	‘sparrow’
e.	*tɕa313*	脚	‘foot’
f.	*dzaŋ24*	近	‘near’

Affricates are produced with friction noise added to the release portion of an obstruent, and thus resemble both stops and fricatives to some extent. Wenzhou differentiates two places of articulation for affricates: alveolar [tsʰa, tsa, dza], and alveolo-palatal [tɕʰ, tɕ, dʑ]. With respect to onset voicing, the affricates of Wenzhou display the same three-member distinction that has been established for the plosives, and the same register division holds (low register tones with “voiced” affricate onset syllables, high register tones with voiceless and voiceless aspirated onset syllables).

The alveolo-palatal affricates [tʰɕ, tɕ, dʑ], like their fricative counterparts [ɕ, ʑ], induce a palatalization of the initial part of the following nucleus vowel if it is [–high], as demonstrated by the narrow transcriptions of the examples in (3): in (3a-c), in which the nucleus contains a [+high] vowel as first or only vowel ([ɵ], [u], and [i], respectively), no palatalized coarticulation can be found. In contrast, the palatal place of articulation of the onset affricates in (3d-h) is spread onto the beginning of the vowel articulation of the [–high] vowels.

(3)	Citation forms	Hanzi	Translation	Narrow transcription
a.	*tɕø33*	专	‘special’	[tɕø]
b.	*tɕuɔ33*	桩	‘stake, pile’	[tɕuɔ]
c.	*tɕiɛ33*	招	‘to attract’	[tɕiɛ]
d.	*tɕa313*	穿	‘to wear’	[tɕʲa]
e.	*tɕau42*	救	‘to rescue’	[tɕʲau]
f.	*tɕou33*	州	‘prefecture’	[tɕʲou]
g.	*tɕaŋ33*	金	‘gold’	[tɕʲaŋ]
h.	*tɕoŋ33*	中	‘middle; in’	[tɕʲoŋ]

In this sketch, the phenomenon will be regarded as regular onset-vowel coarticulation process for the palatal affricates and fricatives, and consequently not be indicated in the broad transcription. In traditional descriptions of other Chinese dialects, the palatal coarticulation between onsets and non-high vowels is sometimes transcribed with a triphthong starting with an [i] (e.g. [ɕiou], [ɕiaŋ] in the phonetic description of Standard Chinese in Lee & Zee 2003). However,

the duration of the palatal coarticulation in Wenzhou is very short, and only lasts less than 50 ms into the articulation of the vowel, as illustrated in Figure 2.2 For this reason, it will be treated as coarticulation here.

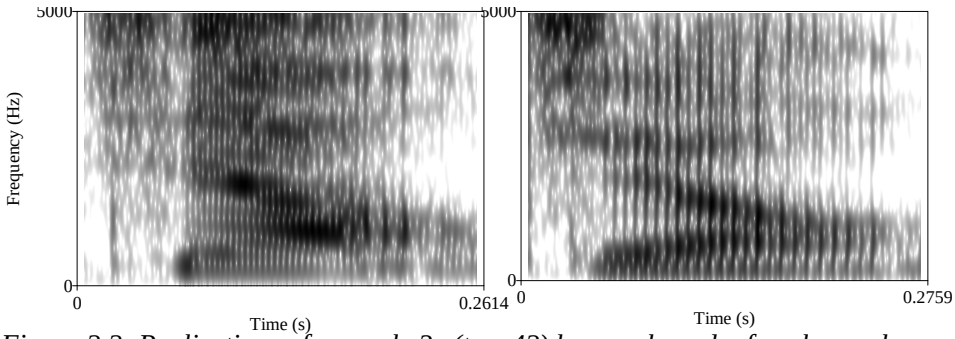


Figure 2.2: Realizations of example 3e (tɕau42) by a male and a female speaker.

2.2.1.3 Nasals

(4)	Citation forms	Hanzi	Translation
a.	*ma24*	买	'to buy'
b.	*na24*	乃	'knife'
c.	*ɲa24*	鸟	'bird'
d.	*ŋa24*	眼	'pupil of the eye'

The nasal series of Wenzhou contains four different places of articulation: bilabial [m], alveolar [n], palatal [ɲ], and velar [ŋ]. Traditionally, nasals are counted as sonorous onsets, and should therefore be expected to pattern with voiced obstruents and only co-occur with low register tones. However, according to the descriptions in Zhengzhang 2008 and the observations from the recordings with the young speakers, all four nasals can co-occur with both high and low register tones. An example for this is given in the minimal pair in (5k-l), and illustrated in Figure 2.3: the syllable [mi] with low and high level tone, as uttered by the same female speaker. The average F_0 in the two examples is ~220 Hz for the low level (5k), and ~245 Hz for the high level (5l).

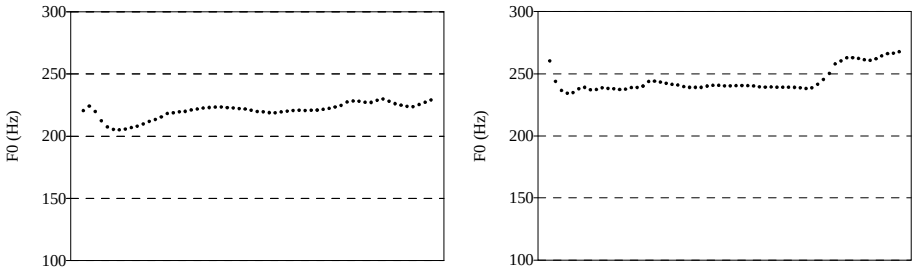


Figure 2.3: Realizations of example 5k (mi11, mean F_0 = 220 Hz) and 5l (mi33, mean F_0 = 245 Hz) by the same female speaker.

The velar nasal [ŋ] deserves special mentioning because of its phonotactic behavior: it can occur as onset consonant before monophthongs and diphthongs (5a-c), as coda consonant following the non-high vowels [a], [ɜ] and [o] (5d-f), and it can stand on its own as a syllabic nasal and carry different tones (5g-j).

(5)	Citation forms	Hanzi	Translation
a.	*ŋa31*	崖	‘cliff’
b.	*ŋu31*	牙	‘tooth’
c.	*ŋau31*	牛	‘ox’
d.	*maŋ31*	门	‘door’
e.	*mɜŋ31*	明	‘bright, clear’
f.	*moŋ11*	梦	‘dream’
g.	*ŋ11*	卧	‘to lie down’
h.	*ŋ24*	我	‘I’
i.	*ŋ24*	五	‘five’
j.	*ŋ31*	儿	‘child’
k.	*mi11*	面	‘face’
l.	*mi33*	眯	‘to narrow one's eyes’

In connected speech, such as phrases and sentences, words which only consist of a syllabified nasal can be merged with preceding (open) syllables and then resemble coda consonants (see also Pan 1991). Thereby, the actual syllable

boundaries can be obscured, which has consequences for the phonotactic analysis (see section 2.3.1 for details).

2.2.1.4 Fricatives

(6)	Citation forms	Hanzi	Translation
a.	*fa42*	泛	'flood'
b.	*va31*	凡	'ordinary'
c.	*sa42*	帅	'graceful'
d.	*za31*	柴	'firewood'
e.	*ɕa35*	晓	'dawn'
f.	*ha42*	喝	'to drink'
g.	*ɦa31*	鞋	'shoe'
h.	*ʋa33*	歪	'crooked'
i.	*va33*	弯	'curve, bend'

Fricatives in Wenzhou differentiate four different places of articulation: labiodental [f, v], alveolar [s, z], alveolo-palatal [ɕ], and glottal [h, ɦ]. Like for the plosives and affricates, there is a correlation between tonal register and onset voicing of the fricatives. As noted in Qian 1992, the bilabial fricatives are subject to ongoing language change in many Wu dialects, in the sense that younger speakers seem to use bilabial [f] and [v] more consistently, whereas older speakers vary between [v] and [β] and between [f] and [ɸ]. In the recordings with the young speakers here, the only noticeable variation occurred for [v], which was sometimes produced in a more approximant-like manner as [ʋ]. Since no phonemic alternation between an approximant [ʋ] and a fricative [v] has been reported in any previous description of Wenzhou, it is assumed here that the approximant is a free variation of the pronunciation of the phoneme [v]. The difference is illustrated with the two examples (6h) and (6i).

The alveolo-palatal fricative [ɕ] is the only fricative with no voiced counterpart in the fricative series of Wenzhou. It only co-occurs with high register tones, and can therefore be deemed "voiceless". Comparable to the alveolo-palatal affricates, [ɕ]-initial syllables induce palatalization of the beginning of the following vowel if it is [–high]. Narrow transcriptions of some examples in (7) illustrate this process, but in accordance with the discussion in the affricate section, the broad transcription in the remainder of this thesis will not reflect this distinction.

(7)	Citation forms	Hanzi	Translation	Narrow transcription
a.	*ɕø42*	汉	‘Chinese’	[ɕø]
b.	*ɕiɛ33*	稍	‘a little’	[ɕiɛ]
c.	*ɕuo33*	霜	‘frost’	[ɕuɔ]
d.	*ɕa35*	晓	‘dawn’	[ɕʲa]
e.	*ɕou33*	收	‘to gather’	[ɕʲou]

From a cross-linguistic point of view, the “voiced” glottal fricative [ɦ] is rare in many other language families of the world, but very common in the subgroup of the Wu dialects.

2.2.1.5 Approximants

(8)	Citation forms	Hanzi	Translation
a.	*ja212*	药	‘medicine’
d.	*la212*	兰	‘orchid’
c.	*la33*	拉	‘pull, drag’
d.	*la11*	烂	‘messy’

Like the nasals, the approximant onset consonants in Wenzhou can co-occur with both high and low register tones. Figure 2.4 illustrates the realizations of the examples in (8c) and (8d) by the same female speaker. The syllable [la] with high level tone (8c) has an average F_0 of ~214 Hz, while the low level example (8d) has an average F_0 of ~188 Hz.

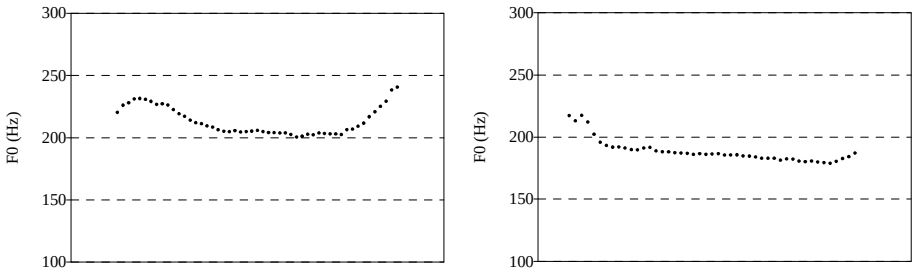


Figure 2.4: Realizations of example 8c (la33, mean F_0 = 214 Hz, left) and 8d (la11, mean F_0 = 188 Hz, right) by a female speaker.

2.2.2 Vowels

2.2.2.1 Monophthongs

Due to numerous co-occurrence constraints and some gaps in the syllable inventory, the velar plosives [g, k, k^h] are the only onsets that can co-occur with all monophthongs in Wenzhou. In (9), a series of (near) minimal pairs is given. The overview of possible Wenzhou syllables in Zhengzhang 2008 contains no example of [kɵ], which is why the nearest approximation [k^hɵ] is listed in (9).

(9)	Citation forms	Hanzi	Translation
a.	*ka35*	解	'to separate, explain'
b.	*kɜ35*	稿	'draft, sketch'
c.	*k^hɵ35*	砍	'to chop, hack'
d.	*ko35*	假	'false'
e.	*ku35*	果	'fruit'
f.	*kø35*	赶	'to catch up'
g.	*ki35*	几	'small table'

The quality of some of the vowels in Wenzhou is different from the corresponding vowels in Mandarin dialects. Unlike the [y]-vowel in the northern dialects for example, the vowel that is transcribed as [ø] here and as [y] in most descriptions of Wenzhou (Hou 1998; You 2002; You & Yang 1998; Zhengzhang 2008) is retracted, less rounded, and also lower in quality than its cognate [y] in Standard Chinese. Conversely, the sound that is transcribed as [ø] or [$ø^y$] in Hou 1998; You & Yang 1998; Zhengzhang 2008 is retracted for the young speakers recorded here and articulated in a central position, and will consequently be transcribed here as [ɵ]. For the vowel [ɜ], only one symbol will be used in this sketch, whereas the descriptions in Hou 1998; You 2002; You & Yang 1998; Zhengzhang 2008 use three different symbols ([ɜ], [ɛ], and either [e] or [ə]) to represent the phonetic value of the sound in different contexts.

It has been observed for a number of Wu dialects that the back high vowel [u] tends to have a tendency to be unrounded (Pan 1991), and that specifically in the Wenzhou phoneme inventory, "[u] has developed into /ʋ/ in the Wenzhou area, a vocalized version of /v/." (ibid: 260). For some of the Wenzhou speakers recorded here, there is indeed a tendency for less lip protrusion, so that the lips are more spread during the vowel production. This can co-occur with some (optional) lip trilling or breathing during the vowel

production, so that the vowel indeed assumes a more fricative-like timbre. Also, as in many other Wu dialects, the two phonemes [u] and [o] are closer together than in other languages, largely through a raising of [o]. The two phonemes remain distinguishable by native listeners, but may fall in the same phoneme category for listeners from Indo-European languages such as German or Dutch.

As in many other Chinese dialects, Wenzhou has a regular alternation for the sound that occurs after alveolar fricative and affricate onset consonants in complementary distribution with [i] in other contexts. This sound is traditionally transcribed with the <ɿ>-grapheme in Chinese linguistics, and is transcribed here as a syllabic fricative [z̩]. Phonetically, the tongue position of the onset fricatives is preserved throughout the pronunciation of the syllable, and merely the voicing is changed at the transition from onset to nucleus in the case of voiceless onset consonants (Ladefoged & Maddieson 1996). The transcriptions in (10) serve to illustrate the process, but this sketch remains agnostic in the debate on the "correct" representation of the "fricative vowels" of Chinese. For discussion, see e.g. Zhang 2006: 53-57 and Duanmu 2007.

(10)	Citation forms	Hanzi	Translation
a.	*ts^{h}z̩42*	处	'place'
b.	*tsz̩42*	至	'to, until'
c.	*dz̩31*	迟	'tardy'
d.	*sz̩42*	似	'similar'
e.	*zz̩31*	视	'to look'

In contrast to Mandarin, which allows a wide range of monophthongs and diphthongs before the nasal codas [n] and [ŋ], Wenzhou only allows the non-high vowels /a/, /ɜ/, and /o/ before the velar nasal [ŋ], which is also the only possible coda consonant in the dialect. In the combination with this coda consonant, the realization of /ɜ/ and /o/ is centralized for all three vowels, compared to the monophthong realization in open syllables. This is reminiscent of the difference between vowels occurring in open and closed syllables in Shanghai Chinese, where the vowels exhibit a reduced acoustic vowel space in closed syllables (Chen 2008). A narrow transcription of the examples in (11) would transcribe them as [mɐŋ], [məŋ], and [moŋ], with the latter being lower and more front than the [o] in open syllables.

(11)	Citation forms	Hanzi	Translation
a.	*maŋ11*	闷	'stuffy'
b.	*mɜŋ11*	命	'life'
c.	*moŋ11*	梦	'dream'

2.2.2.2 Diphthongs

(12)	Citation forms	Hanzi	Translation
a.	*t^hai33*	推	'to push'
b.	*t^hɛi33*	梯	'ladder'
c.	*t^hau33*	偷	'to steal'
d.	*t^hou33*	拖	'to pull, drag'
e.	*t^hiɛ33*	挑	'to choose'
f.	*t^huɔ33*	汤	'soup'

As can be seen in (12), when counting [ɵ^{y}] as an alternative realization of the monophthong [ɵ], Wenzhou differentiates six diphthongs. Most Chinese grammars further differentiate combinations of the diphthongs with an initial [i] or [ø], thus deriving additional diphthongs with onglides or triphthongs. In the present investigation, they will be regarded as regular allophones, in accordance with the above analysis of the variants of [ɕ] inducing palatalization in [–high] vowels (cf. section 2.2.1.4, examples in (7)).

2.3 Phonotactics

2.3.1 Well-formed syllable structure

Since the velar nasal coda in Wenzhou can only co-occur with the monophthong vowels [a], [ɛ] or [ɔ], the maximal syllable in Wenzhou is *CV*(*X*), with *X* being either a second vowel (in diphthongs) or a velar nasal. Classifying Wenzhou as having an obligatory onset presupposes the analysis of the 0-onset as glottal stop (see section 2.2.1.1 for discussion), and the analysis of affricates as single *C*. Furthermore, as discussed in section 2.2.1.3, the velar nasal can be syllabic, and obligatorily be combined with *CVV* syllables. Unlike the [ɚ]suffix in Beijing Mandarin and other Mandarin dialects though (Duanmu 2007), the [ŋ] in Wenzhou does not merge with its host syllables, but is merely added to their

existing structure. This may create superheavy *CVVN* syllables, as in example (13).

(13)	Citation forms	Hanzi	Translation
a.	*muɔ33*	猫	'cat'
b.	*muɔŋ31*	猫儿	'cat'

2.3.2 Attested combinations of consonants and vowels

In line with other Chinese dialects, Wenzhou displays certain co-occurrence restrictions when it comes to the combination of the phonemes of the language into valid syllables. Some of these have already been covered, such as the allophony on [+high, +front] vowels following alveolar fricatives and aspirates (section 2.2.2.1), and the restriction of the coda consonant [ŋ] to only co-occur with allophones of three of the monophthong vowels in single syllables (sections 2.2.2.1 and 2.3.1). Outside of these, the only consistent co-occurrence restriction that could be found across all possible onset-rime combinations is a restriction for alveolo-palatal onsets to occur before rimes that start with mid vowels ([ɜ], [ɵ], [ɛi], [ɜŋ].)

However, this does not mean that all other possible onset-rime combinations indeed occur in Wenzhou. Appendix 2.2 gives an overview of attested syllables, and illustrates some mismatches between the syllables listed in the most recent comprehensive description of Wenzhou (Zhengzhang 2008) and the recordings with the young speakers for this thesis. These mismatches occurred in both directions, with some speakers realizing characters with syllables that are not listed in Zhengzhang 2008, and with speakers reporting in elicitation sessions that they could not think of a word in Wenzhou that would be pronounced with a particular onset-rime combination that is listed as possible syllable in Zhengzhang 2008. It is unclear whether these mismatches constitute mispronunciations on the part of the speakers, or whether they represent ongoing language change. For cases where no word with a particular onset-rime combination could be elicited, but where the speaker did not deny the existence of the particular combination, the syllable is listed in appendix 2.2 as existent.

2.4 Tone inventory in monosyllables

(14)	Syllable	Tone type[3]	MC class	Hanzi	Translation
a.	*ta33*	high level	Ia	单	'alone, single'
b.	*ta35*	high rising	IIa	胆	'guts'
c.	*ta42*	high falling	IIIa	带	'belt, zone'
d.	*ta313*	high dipping	IVa	搭	'to put up, hang over'
e.	*da11*	low level	IIIb	潭	'pond'
f.	*da24*	low rising	IIb	淡	'thin, light'
g.	*da31*	low falling	Ib	谈	'to chat'
h.	*da212*	low dipping	IVb	踏	'to step on sth.'

The examples in (14) above illustrate the eight citation tones of Wenzhou. As mentioned in the description of plosive onsets in Wenzhou, the eight tones of Wenzhou are commonly split into two registers, each containing four tones. The register division aligns to some extent with the "voicing" properties of the onset consonant, but some onsets, such as nasals and approximants, can occur with both high and low register tones. It has been proposed that the register division arose from a spread of the voicing properties of the onset onto the vowel, and that the pitch differentiation occurred as a secondary process via tonogenesis (Duanmu 1992). However, the connection between the register of a tone and its height is non-automatic, and in some Chinese dialects, "high" register tones may in fact be phonetically lower than "low" register tones (Chao 1967, see also Hayward et al. 2003).

For the Wenzhou "high" and "low" register tones, a difference in tone height can be confirmed at least for the level and falling tones, so the traditional terminology will be kept for ease of reference. However, as the description of the acoustic properties of the tones shows, it is overly simplistic to think of the eight tones as four contours being realized in two different areas of the speakers'

[3] In order to facilitate reference to the eight tones in Wenzhou, they are described using a combination of the register (high/ low) and the tone type (level, rising, falling, dipping), which are commonly used translations of the Middle Chinese tone classes (e.g. in Rose 2001 and Chen 2000, the latter of who substitutes "even" for "level"). However, the terms should not be taken to be descriptive of the tonal trajectories: As already noted in Rose 2008, the "dipping" tones of the young speakers largely lack the final rise, and realize the "dipping" tones with a falling contour instead.

pitch ranges, as the terminology might suggest. For this reason, the properties of the tones will be described in as much detail as possible.

In traditional Chinese descriptions, tones are often transcribed with the help of the number system after Chao (1948, see Duanmu 2007 for critical discussion). However, the exact Chao number values that are used to describe the tones in a dialect may vary between transcribers, for example whether e.g. a falling tone is described as 53 or 42, and whether tone duration is reflected in the number system (as for example in the description in Rose 2000). Table 2.2 lists some Chao number descriptions of Wenzhou tones that have been published in the recent literature, before describing in detail the tonal realization observed in the young speakers who were recorded for this phonology sketch.

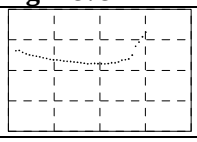
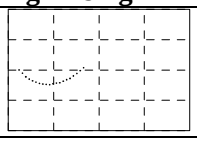
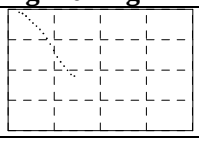
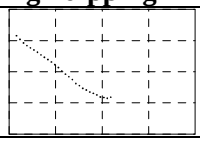
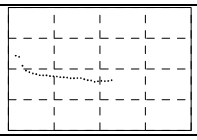
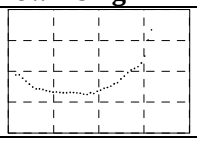
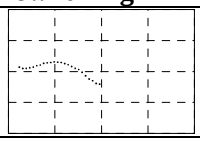
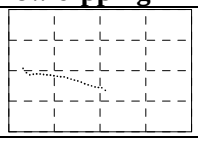
Table 2.2: Numeric descriptions of Wenzhou citation tones in the recent literature.

	High level	**High rising**	**High falling**	**High dipping**	**Low level**	**Low rising**	**Low falling**	**Low dipping**
Chen 2000	M (33)	MH (35)	HM (53)	Lq (313)	L (11)	MH (35)	ML (31)	Lq (212)
Nakajima 1983	44	45	42	323	31	23	242	212
Norman 1988	44	45	42	23	11	24	31	12
Hou 1998	33	35	42	213	22	35	31	213
Qian 1992	44	35	52	423	22	24	31	323
Rose 2000	33	34	52	3312	222	114	331	2212
You 2002; You & Yang 1998	33	35	42	313	11	24	31	212
Zhengzhang 1964b	44	45	42	323	22	34	31	212
Zhengzhang 2008	33	45	42	313	11	34	31	212

In order to illustrate the tonal contours of the young Wenzhou speakers, Table 2.3 depicts the F_0 trajectories of the examples in (15) as uttered by a female speaker. In order to allow for a comparison not only of the F_0 contour and height, but also of the durational properties of the tones, the pitch tracks are depicted

initially in a window of 400 ms length, with the left edge of window being aligned with the plosive burst of the onset consonant.

Table 2.3: Representative F_0 tracks of the eight lexical tones of Wenzhou on alveolar plosive onset syllables (examples 15a-h), recorded with a female speaker. F_0 range 100-300 Hz, intermediate lines = 150, 200, 250 Hz and 0.1, 0.2, and 0.3 seconds.

High level	**High rising**	**High falling**	**High dipping**
			
Low level	**Low rising**	**Low falling**	**Low dipping**
			

Visual and auditory inspection of the lexical tones of Wenzhou yield the following observations.

(15) **Observations on lexical tones in young Wenzhou speakers**

- The level tones are both situated in the middle of the speakers' pitch ranges, with notably the "high" level tone having a rather "mid" quality. Especially for the male speakers, the resulting shortage of distinctiveness is often compensated by lengthening of the low level tone. Both tones display a declining tendency, but for the female speakers, the high level tone may be accompanied by a short rising offset.
- Both rising tones are characterized by falling-level trajectories through the first half of the tonal duration. The "high" and "low" rising tones are often realized on similar F_0 height, which led e.g. Rose (2008) to claim that in the young speakers, the two tones have merged. However, in the young speakers recorded here, they are often distinguished by lengthening of the low register tone.
- For the falling tones, the distinction between "high" and "low" register is most clearly visible in the beginning of the tonal contour, which is

characterized by a short rising portion before falling to a comparable tone height as the high falling tone. Both tones are rather short of duration, most notably the high falling tone.

- For the dipping tones, the most striking observation is the lack of the final rise of the "dipping" contour. As the falling tones, the two dipping tones are distinguished mainly by the onset height of the fall, which is lower for the low-register tones.

As can be seen from the above description, several factors have to be taken into account when characterizing the tone contours of Wenzhou. Besides the actual tonal directions (which do not suffice any more to distinguish the falling from the dipping tones), the length of the tones and the presence/absence of the "depressor" effect at the onset of the tone are important for the distinction between "high" and "low" register tones.

2.5 Tone sandhi

2.5.1 Terminology

The term "tone sandhi" has been used in the literature with reference to two different phenomena: tonal coarticulation and rule-based tonal change (Chen 2000). The former phenomena include processes as described for example in Xu 1997 for Standard Chinese, involving assimilatory/dissimilatory and transitory processes that take place between two adjacent tones when they are produced in the context of one another. On the other hand, the term "tone sandhi" has also been used to refer to tone change processes in which at least one of the involved tones changes its phonological shape in a categorical manner. A famous example of such a change is the so-called "third tone sandhi" of Beijing Mandarin, in which, according to the common analysis, a third tone followed by another third tone within a specific prosodic domain is realized as a second tone (for discussions pro/contra full phonological neutralization see e.g. Chen & Yuan 2007; Peng 2000; Xu 1997; Zhang & Lai 2006, 2010).

These two uses of the term "tone sandhi" have been conflated on the basis of the argument that the two processes are not fundamentally different, and that they should be treated as gradual variants of each other (Chen 2000). However, as will become clear in the remainder of this thesis, for Wenzhou, it clearly makes more sense to keep the two terms apart. Therefore, the term "tone sandhi" will be used in this thesis to refer exclusively to such tone change phenomena in which, as stated above, at least one tone changes its phonological

shape in a categorical and rule-based manner. This implies that (i) the tonal change cannot be entirely explained as articulatorily conditioned modification of the shape of either of the lexical tones, and (ii) that the process is stable over different segmental TBUs, as long as they share similar morphosyntactic or prosodic conditioning environment.

While "tone sandhi" and "tonal coarticulation" are thus differentiated in terminology, both processes undoubtedly may serve the same function, namely simplification of articulation. Often, phonological accounts of tone sandhi processes strive to lend them motivation by referencing how they promote ease of articulation, or contribute to the reduction of markedness. Therefore, the above differentiation should be regarded as distinction of definition, but it does not exclude the possibility of functional overlap.

2.5.2 Domain of application of tone sandhi in Wenzhou Chinese

Commonly, the term "tone sandhi" with respect to Wenzhou Chinese describes processes targeting lexical words which consist of more than one syllable. The most common tone sandhi cited in the literature involves disyllabic lexical compound nouns, verbs, and adjectives, and it has also been argued to apply to disyllabic modifier constructions (Chen 2000). In the recordings of the young speakers, this pattern could be confirmed, as the speakers realized disyllabic lexical compound nouns, verbs, and adjectives with regular disyllabic contours, as described in the following section. The same was true for some adjective-noun-collocations. In contrast, in verb-object-collocations with two monosyllabic lexical items, the most prevalent pattern was the adoption of a phrasal prosody that retained the tonal value of the strong prosodic position (non-head) and reduced the tonal value of the prosodically non-strong position to low level. For an in-depth analysis of the domain of application of disyllabic tone sandhi, the reader is referred to Chapter 3 of this dissertation.

2.5.3 Patterns of tone sandhi contours in Wenzhou Chinese

A description of the patterns of (disyllabic) tone sandhi in Wenzhou has represented a challenge for phonologists, since the observed patterns seem to be "[d]efying explanation" (Rose 2004, citing Chen 2000). What is meant by that is that the connection between the citation forms and the sandhi contours in connected structures do not seem to be easily explicable through standard phonological processes such as spreading/shifting of tones, substitution, or neutralization (Chen 2000). Even if some of the patterns can be explained in a

rule-based manner, no unifying explanation has been proposed which could account for all observed tone sandhi contours (see also Yue-Hashimoto 1987).

For the current description, the matter is further complicated by the observation that the tone sandhi contours that were recorded with the young speakers do not fully match the tone sandhi contours that have been described even in the most recent literature (e.g. Chen 2000; Rose 2000, 2001, 2002, 2004; You 2002). This is caused partially by mismatches in the realization of the citation forms of the individual syllables, some of which are realized by the young speakers in a way that differs from previous descriptions on a segmental and/or on a tonal level. On the other hand, the observed tone sandhi contours are also clearly different even for those disyllabic examples where the realizations of the citation tones match published descriptions. In appendix 2.1, an overview is given of differences between some of the published descriptions and the tone sandhi contours recorded for this thesis.

Table 2.4 gives an overview of the tone sandhi contours that were observed in the young speakers, based on a broad transcription in Chao numbers.

Table 2.4: Numerical overview of tone sandhi contours observed in the speech of the young Wenzhou speakers.

<table>
<tr><th rowspan="2">First σ</th><th colspan="8">Second σ</th></tr>
<tr><th>33</th><th>31</th><th>35</th><th>24</th><th>42</th><th>11</th><th>313</th><th>212</th></tr>
<tr><td>33</td><td colspan="3">42.31</td><td>44.22</td><td rowspan="2">22.44</td><td rowspan="2">22.35</td><td rowspan="8" colspan="2">35.42</td></tr>
<tr><td>31</td><td colspan="2">31.22</td><td colspan="2">35.22</td></tr>
<tr><td>35</td><td>44.22</td><td>42.31</td><td colspan="2">44.22</td><td>42.31</td><td>44.22</td></tr>
<tr><td>24</td><td colspan="2">31.22</td><td colspan="2">35.22</td><td colspan="2">31.22</td></tr>
<tr><td>42</td><td>44.22</td><td>42.31</td><td colspan="2">44.22</td><td>42.31</td><td>44.22</td></tr>
<tr><td>11</td><td>35.22</td><td>31.22</td><td colspan="2">35.22</td><td>31.22</td><td>35.22</td></tr>
<tr><td>313</td><td rowspan="2">22.33</td><td rowspan="2">42.31</td><td rowspan="2" colspan="2">22.35</td><td rowspan="2">22.42</td><td rowspan="2">42.31</td></tr>
<tr><td>212</td></tr>
</table>

As can be seen from Table 2.4, even though the tone sandhi contours are different from the descriptions in the literature, there is still some neutralization in the patterns, in that specific tonal combinations result in tone sandhi contours that are acoustically comparable to the tone sandhi contours of other tonal

combinations. In (16), a verbal description of these tone sandhi contours and the tonal combinations they result from is listed.[4] The individual patterns are described in greater detail in the corresponding sections below.

(16) **Types of tone sandhi contours**
a. Rise – fall contour
→ x+dipping
b. Fall – low level contour
→ x+31, rising+42, 42/11+42, 33+33, 31+33, 33+35, 24+33, 24+11, dipping+11
c. Rise – low level contour
→ 31+rising, 24+rising, 11+rising, 11+33
d. Low level – rise contour
→ dipping+rising, 33+11, 31+11
e. High level – low level contour
→ 33+24, 35+rising, 42+rising, 35+33, 35+11, 42+33, 42+11
f. Low level – high level contour
→ 33+42, 31+42, dipping+33
g. Low level – fall contour
→ dipping+42

2.5.3.1 Rise – fall contour (16a)

As can be seen in Table 2.4, the category of dipping tones (212, 313) is the only tone category that displays regularity when appearing on the second syllable in a tone sandhi context. Combined with any other tone on the first syllable, the dipping tones result in a rise-fall tone sandhi contour (= pattern (16a, Figure 2.5).

[4] To differentiate the combinations of citation tones on the individual syllables from the disyllabic tone sandhi contours, the former will be indicated with a plus sign combining the citation tone values of the individual syllables, while the latter will be described with a dot between the tone sandhi contour values of the two syllables.

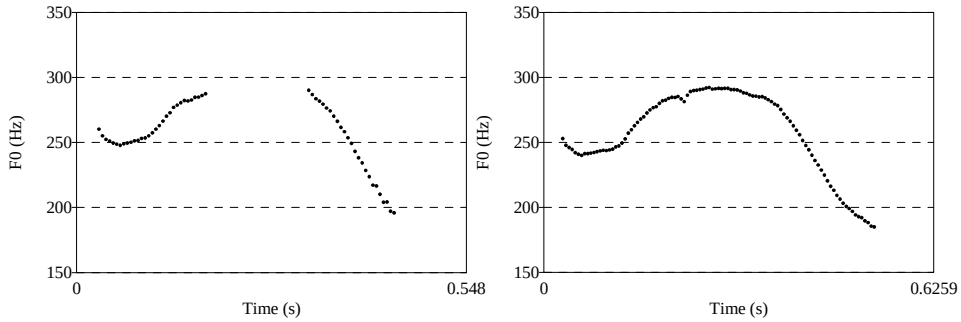


Figure 2.5: Representative pitch tracks of the rise-fall tone sandhi contour with a high register dipping tone (left) or a low register dipping tone (right) on the second syllable. Citation tones left = 35+313, right = 35+212.

In the tone sandhi contour resulting from x+313 combinations, the F_0 contour of the first syllable is characterized by a rise that starts about mid-level of the pitch range of the speaker and remains level for a short period of time, before rising to a high pitch point. The second syllable starts approximately at that point, falls sharply and straight throughout the second syllable, and ends at a pitch point that lies lower than the pitch onset point of the rise on the first syllable. The tone sandhi contours that result from x+212 combinations are characterized by a continuous transition from the rising to the falling part of the contour. Other than that, the pitch trajectory of the x+212 combinations mirrors that of the x+313 combinations, in that the pitch onset of the rise on the first syllable is higher than the pitch offset of the fall on the second syllable.

2.5.3.2 Fall – low level contour (16b)

A second very common tone sandhi contour in the speech of the young Wenzhou speakers involves a falling tone on the first syllable and a low level or declining tone on the second syllable (= pattern 16b, Figure 2.6).

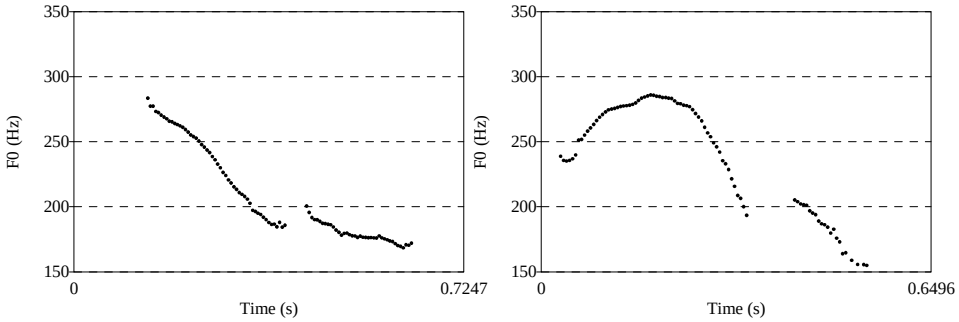


Figure 2.6: Representative pitch tracks of the fall-low level tone sandhi contour with a high register tone (left) or a low register tone (right) on the first syllable. Citation tones left = 33+31, right = 24+31.

Depending on the register on the first syllable, the onset of the fall on the first syllable will be either continuously falling (high register) or rising-falling (low register). In that sense, the shape of the falling tone contour on the first syllable is very similar to the lexically high and low falling tones, which are characterized by a similar difference in tonal direction (cf. Table 2.3 in section 2.4.1). This tone sandhi pattern occurs on all combinations with a low falling tone on the second syllable, and on four combinations with a high falling tone on the second syllable. Outside of these combinations of citation tones, however, the pattern also results in some combinations in which the citation tone on neither syllable is falling. It is therefore difficult to relate this tone sandhi contour to the citation tones on the individual syllables in a transparent way.

2.5.3.3 Rise – low level contour (16c)

Another frequent tone sandhi contour consists of a rising tone on the first syllable, followed by a low level or declining tone on the second syllable (= pattern 16c, Figure 2.7).

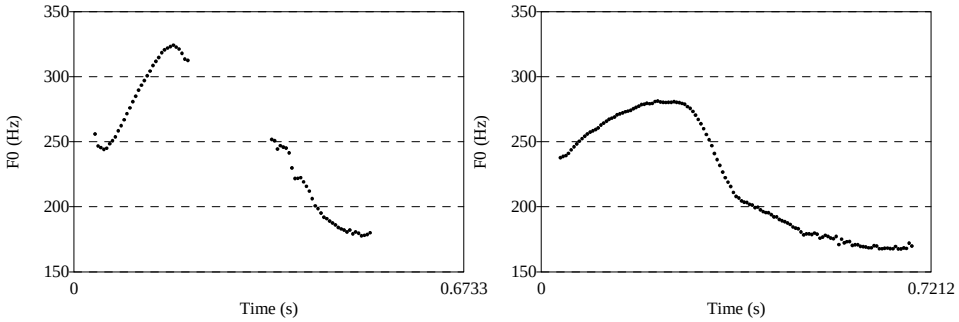


Figure 2.7: Representative pitch tracks of the rise-low level tone sandhi contour with a high register tone (left) or a low register tone (right) on the second syllable. Citation tones left = 11+35, right = 31+24.

As with the previous pattern, it can result from a combination of a syllable with a rising citation tone in second position, but since this tone sandhi pattern is also found on the combination 11+33, again no straightforward generalization can be made. This tone sandhi pattern shows some resemblance with the depressed falling-low level pattern described above. However, as can be seen from a comparison between Figures 2.6 and 2.7, the turning point in the connected falling-low level pattern is earlier than that in the connected rising-low level pattern. This suggests that the two patterns are not variants of each other, but indeed consist of different tonal contours.

2.5.3.4 Low level – rise contour (16d)

The inverse pattern, with a rise on the second instead of the first syllable, is found on the combination of dipping tones on the first syllable with rising tones on the second syllable, as well as a couple of other citation tone combinations (= pattern 16d, Figure 2.8).

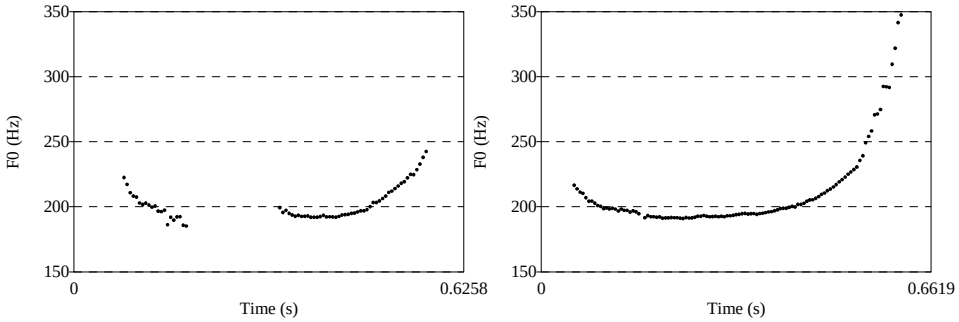


Figure 2.8: Representative pitch tracks of the low level-rise tone sandhi contour with a high register tone (left) or a low register tone (right) on the second syllable. Citation tones left = 313+35, right = 313+24.

2.5.3.5 High level – low level contour (16e)

Another tone sandhi contour that can be found across different combinations of citation tones is the combination of a high and a low level tone, mostly as result of a combination of two contour citation tones (= pattern 16e, Figure 2.9).

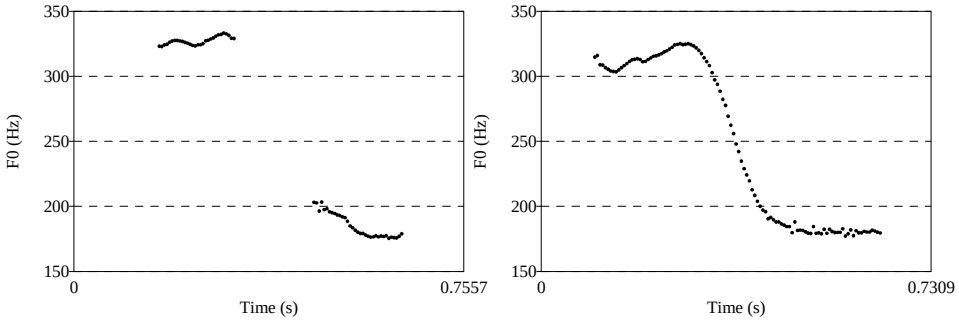


Figure 2.9: Representative pitch tracks of the high level-low level tone sandhi contour with a high register tone (left) or a low register tone (right) on the second syllable. Citation tones left = 35+35, right = 33+24.

One of the speakers of this sketch classified the high level tone of this tone sandhi contour, when played in isolation, as a high falling tone. This is not surprising, since it possesses two characteristic aspects of the latter, namely the shortness and the high F_0 level. In contrast, as discussed in section 2.4.1, the so-

called "high" level lexical tone of Wenzhou is situated in the middle of the speakers' pitch ranges, and lower than the high level tone in the sandhi contour discussed here.

2.5.3.6 Low level – high level contour (16f)

The inverse pattern to the just mentioned tone sandhi contour, namely a low combined with a high level tone, is only found in two combinations of citation tones, namely the high level or low falling tone on the first syllable plus the high falling tone (42) on the second syllable (= pattern 16f, Figure 2.10).

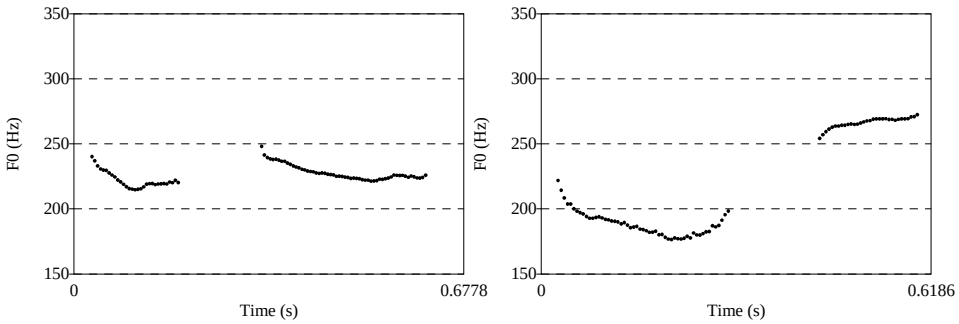


Figure 2.10: Representative pitch tracks of the low level-high level tone sandhi contour with a high register tone (left) or a low register tone (right) on the first syllable. Citation tones left = 313+33, right = 31+42.

As in the previous tone sandhi contour, the pitch trajectory of both syllables is level, but the high level tone on the second syllable does not display the extra-high F_0 target that it did in the previous pattern. Rather, it is situated in the mid range of the speaker's F_0 range, similar to the lexical "high" level tone.

Within the tone sandhi pattern (16f), two types of combinations of citation tones pattern together: 33/31+42, and dipping+33. While they are grouped into the same contour type in this analysis, an important distinction can still be made between these two types of citation tone combinations: the resulting length distribution in the tone sandhi contours is different. This can also be seen in Figure 2.8: the combinations of dipping+33 result in a short-long tone sandhi contour, while the combinations 33/31+42 result in a long-short tone sandhi contour. A possible explanation for this can be found in the underlyingly short duration of the high falling tone (42), which might be preserved in sandhi

contours even when the originally falling contour is replaced by a different tonal movement.

2.5.3.7 Low level – fall contour (16g)

The last tone sandhi contour results from the combination of a dipping tone on the first syllable and the high falling tone on the second syllable, and turns out as low level-high falling (= pattern 16g, Figure 2.11).

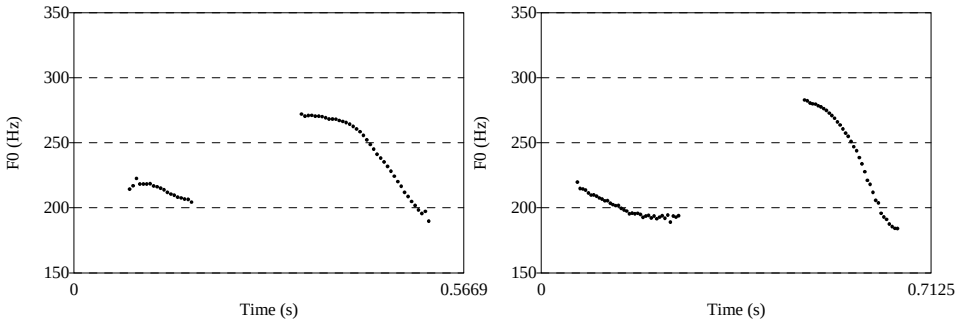


Figure 2.11: Representative pitch tracks of the low level-falling tone sandhi contour with a high register tone (left) or a low register tone (right) on the first syllable. Citation tones left = 313+42, right = 212+42.

2.5.4 Generalizations on the tone sandhi contours in Wenzhou Chinese

In the foregoing section, the tone sandhi patterns were grouped and described according to the similarities in the resulting tone sandhi contours. Abstracting away from the individual contours, this section will focus on the relationship between the citation tones on the individual syllables and the tone sandhi contours in different contexts, and discuss the occurring transparent connections.

When looking only at the combinations with dipping tones on the first syllable, it becomes apparent that the second syllables following the dipping tones mostly preserve their citation forms, while the first syllable often takes on a low level trajectory. This has led to generalizations in the literature in which it is claimed that in combinations of a dipping tone with any other tone, both syllables preserve their citation tone values, and the dipping tone trajectory becomes flattened and shortened. However, with the younger speakers, this generalization only holds true when excluding the patterns in which two dipping

tones combine. These latter patterns are realized with a rise-fall contour (= pattern 16a).

Another research question connected to the above mentioned observation concerns the distribution of "prominence" between the syllables of a word. A common assumption is that in disyllabic words, one syllable is "weak" whereas the other is "strong", and that this imbalance in prominence conditions or restricts the application of tone sandhi. Wenzhou has typically been regarded as a "right-prominent" language (Chen 2000; Yip 1999). This would entail the expectation that in the tone sandhi patterns, the rightmost syllables should generally be longer and carry more complex tonal contours.

As can be seen in the descriptions of the tone sandhi contours above, this generalization does not describe the findings for the young speakers accurately. While it is true that some tone sandhi patterns display an uneven distribution of duration that is independent of the segments of the syllables, these uneven distributions cannot be generalized across the entire data set. Rather, they seem to be connected to certain types of tones. For example, a dipping tone on the first syllable, combined with anything else than a dipping tone on the second syllable, will come out as a short low level tone in the tone sandhi contour. However, the same tone combined with another dipping tone will surface in an tone sandhi contour which is balanced for duration. Similarly, syllables with lexically falling tones can come out in tone sandhi contours as short when appearing either on the first or second syllable, but this is not a necessary process.

A couple of generalizations can be observed in the tone sandhi patterns of the young speakers, which are summarized in (17) and numerically in Table 2.5. The descriptions are verbal and no Chao numbers are given, so that tones with the same shape and different registers can be grouped together more easily.

(17) **Generalizations about tone sandhi patterns in young Wenzhou speakers**

a. Any tone combined with a dipping tone on the second syllable will result in a rise-fall contour.

b. A dipping tone on the first syllable combined with any other tone except another dipping tone on the second syllable will result in a pattern in which the first syllable is short and low level, and the second syllable retains its citation tone. In the case of the low falling and low level tone on the second syllable, the first syllable will become high falling, which results in a continuously falling contour over the two syllables.

c. A combination of any tone on the first syllable and a low falling tone on the second syllable will also always result in a continuously falling contour over the two syllables.

d. A high falling tone on the second syllable will result in a continuously falling contour over the two syllables when combined with any tone except the dipping tones and the tones of Middle Chinese class I (33, 31) on the first syllable.

e. A rising tone on the second syllable combined with any low register tone except the low dipping tone on the first syllable will result in a rise-low level contour.

f. A rising tone on the second syllable combined with the high rising or the high falling tone on the first syllable will result in a high level-low level contour.

g. A low or high level tone on the second syllable combined with the high rising or high falling tone on the first syllable will result in a high level-low level contour.

h. A low rising tone on the first syllable combined with a high or low level tone on the second syllable will result in a continuously falling contour.

i. A low level tone on the first syllable combined with another low level tone or a high level tone on the second syllable will result in a rising-low level tone contour.

j. A high level tone on the second syllable combined with either of the two tones from Middle Chinese class I (33, 31) will result in a continuously falling contour.

k. A high falling tone on the second syllable combined with either of the two tones from Middle Chinese class I (33, 31) will result in a low level-mid level contour.

l. A low level tone on the second syllable combined with either of the two tones from Middle Chinese class I (33, 31) will result in a low level-rising contour.

m. A high level tone on the first syllable will result in a continuously falling contour when combined with a high rising tone on the second syllable, and in a high level-low level contour when combined with a low rising tone on the second syllable.

Table 2.5: Distribution of generalizations on the tone sandhi patterns in the speech of the young Wenzhou speakers.

<table>
<tr><th rowspan="2">First σ</th><th colspan="8">Second σ</th></tr>
<tr><th>33</th><th>31</th><th>35</th><th>24</th><th>42</th><th>11</th><th>313</th><th>212</th></tr>
<tr><td>33</td><td rowspan="2">18j</td><td rowspan="6">18c</td><td colspan="2">18m</td><td rowspan="2">18k</td><td rowspan="2">18l</td><td rowspan="8" colspan="2">18a</td></tr>
<tr><td>31</td><td colspan="2">18e</td></tr>
<tr><td>35</td><td>18g</td><td colspan="2">18f</td><td rowspan="4">18d</td><td>18g</td></tr>
<tr><td>24</td><td>18h</td><td colspan="2">18e</td><td>18h</td></tr>
<tr><td>42</td><td>18g</td><td colspan="2">18f</td><td>18g</td></tr>
<tr><td>11</td><td>18i</td><td colspan="2">18e</td><td>18i</td></tr>
<tr><td>313</td><td rowspan="2" colspan="5">18b</td></tr>
<tr><td>212</td></tr>
</table>

It seems that there is a certain strength ranking between the citation tones that decides which tone in the combination determines the pitch slope of the tone sandhi contour, and which tone can preserve some characteristics of the citation forms. It appears that the dipping tones are strongest in superimposing a uniform contour on the other tones when they are in final position, but weakest when in initial position, where they take on a low level slope while the second syllables retain their citation tones. The only exception to that are the low level and low falling tone, which are disambiguated from the dipping tone on the first syllable by the latter taking on a falling slope.

Further, the rising and falling tones superimpose left-prominent contours on any initial syllable they are combined with, with the exception of class I tones, which behave differently. The rising and falling tones on the second syllable transmit their characteristics (or end points) onto the first syllable, and take on a low level contour themselves. Only when the initial syllable is occupied by a class I tone do the second syllables take on a different tonal contour, either a falling or a high level pitch. The level tones are the odd ones out in most scenarios, and consequently, when they appear in second position, there is much more variability when it comes to the shape of the second tone.