Universiteit
Leiden
The Netherlands

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

Scholz, F.

Citation
Scholz, F. (2012, October 18). *Tone sandhi, prosodic phrasing, and focus marking in Wenzhou Chinese*. *LOT dissertation series*. LOT, Utrecht. Retrieved from https://hdl.handle.net/1887/19983

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/19983

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/19983 holds various files of this Leiden University dissertation.

Author: Scholz, Franziska
Title: Tone sandhi, prosodic phrasing, and focus marking in Wenzhou Chinese
Issue Date: 2012-10-18

Chapter 4

The effect of the tone sandhi domain on focus expression in Wenzhou Chinese

4.1 Introduction

This chapter is concerned with the phonetic marking of contrastive focus, and the influence of prosodic structure on this phonetic marking. For intonation languages, it has been found that unlike other types of focus, contrastive focus can be marked not only on words, but even on individual syllables of words, and (for proficient speakers) even on parts of syllables (Sluijter 1992; Sluijter & van Heuven 1995; van Heuven 1994). In such cases, speakers lend additional prominence to the contrasted parts of the word, using phonetic markers such as F_0/pitch accentuation, lengthening, intensity, and spectral distribution. The examples in (1) illustrate this for English, with the contrasted parts marked in bold and upper case.

(1) a. I said '**COF**fin', not '**MUF**fin'.
b. I said '**P**it', not '**B**it'.
c. I said 'cof**FIN**', not 'cof**FEE**'.

For Dutch, it has been found that in examples comparable to (1c), where the focus contrast is located on a syllable which is lexically unstressed, the pitch accent which normally marks stress is "shifted" to the unstressed syllable by virtue of focus marking. At the same time, the accompanying durational marking of stress is only partially affected by focus: the focused (lexically unstressed) syllable is lengthened, but this lengthening does not fully reverse the durational difference between the two syllables of the word (Sluijter & van Heuven 1995).

These findings suggest that in intonation languages like English, where pitch accentuation serves a dual function both as a primary marker of lexical stress and as a primary marker of focus, focus can trump phonology and determine the location of the pitch accent within a word. At the same time, it appears that duration as a secondary marker for both focus and lexical stress is much less volatile than pitch accentuation: a lexically unstressed syllable is lengthened under focus, but its relative duration in the word is still shorter than

that of a lexically stressed syllable. Apparently, focus cannot fully override the durational marking of stress, so that a residue of durational stress marking remains intact even under contrastive focus on a non-stressed syllable.

The current study looks at Wenzhou Chinese, a tone language with an extensive tonal phonology. In this language (as in many related dialects of the Wu dialect group), disyllabic words represent an important domain for phonological processes, because when two syllables come together as a compound, the lexical tone on both syllables changes in a process called "tone sandhi". In that sense, it can be said that speakers use the tone change to mark for listeners that the two syllables form one compounded constituent, rather than just two adjacent syllables. Therefore, the tone change process serves as a marker for wordhood. This chapter reports the results of an experiment which tests whether contrastive focus can be marked within this tone sandhi domain, at the expense of the phonological marking of wordhood.

4.1.1 Focus marking in Chinese

Across different Chinese dialects and focus domains of differing sizes, two phonetic correlates of focus have been in the center of attention: F_0 and duration modification.[13] However, most of the studies were concerned with the effect of focus on lexical tones in sentences, whereas the current study investigates tones that result from tone sandhi (cf. section 4.1.3). Furthermore, most studies only considered entire lexical words as focus domains, without testing what happens to the focus marking if only a part of a word is focused. The current study aims to combine these two fields of exploration, and investigate focus marking below the word level on sandhi tones.

Of the two phonetic effects that have been identified as correlates of focus in Chinese, F_0 is the more straightforward one: all Chinese dialects are tonal, and differences in the implementation of these tones between focus and non-focus conditions are readily observable in most dialects. Following the detailed exploration of focus effects in Standard Chinese by Xu (1999), F_0 differences between focus and non-focus condition have been reported for several dialects of Chinese, such as Mandarin Chinese (Wang & Xu 2006),

[13] The experiment reported in this chapter specifically investigates the effect of contrastive focus. However, in the discussion of previous findings on focus effects in Chinese, other types of focus, such as focus induced by answering a wh-question, have been included in the literature review. The assumption that both types of focus elicit comparable effects is interesting in itself, but its verification lies outside of the scope of this study.

Cantonese (Gu & Lee 2007b), Shanghai Chinese (Chen 2009), and Taiwan Mandarin (Chen et al. 2009; Xu et al. 2012). Generally, the reported findings can be summarized as F_0 range expansion under focus, such that the F_0 maximum will be higher, and the F_0 minimum will be lower under focus than in the non-focus control conditions.

In most, but not all of these dialects, the on-focus F_0 expansion effect of focus is accompanied by a F_0 compression effect on the post-focal stretch of the target sentence, which can represent an enhancing cue for listeners to determine the location of the focused constituent in the sentence (Chen et al. 2009; Xu et al. 2012). However, taking a closer look at the phonetic nature of "post-focus compression" in Standard Chinese, Chen (2010) reports that the tone contours on post-focal tones differ depending on the preceding tone. Rather than a uniform lowering effect in post-focal condition, different tonal contexts share a weak tonal implementation in post-focal position, which manifests itself in hypoarticulation of the tonal target, a reduced degree of distinctiveness of the tonal contour, and a greater influence of the preceding tones, especially if the latter were in on-focus condition themselves.

While F_0 manipulation may be a widespread phonetic means, it is not a necessary focus marker across all Chinese dialects. As evident in studies on Hong Kong Cantonese (Wu & Xu 2010) and Taiwan Min (Chen et al. 2009; Xu et al. 2012), in some dialects, the implementation of tonal contours is not systematically influenced by the presence or absence of focus. Rather, in these dialects, duration and intensity of focused syllables seem to play a greater role in signaling focus.

Outside of F_0 manipulation, duration manipulation has been reported as the other stable cue to different focus conditions in Chinese dialects. Lengthening of new or contrastively focused syllables has been found in Taiwan Mandarin (Pan et al. 2005), Taiwan Min (Chen et al. 2009; Xu et al. 2012), Cantonese (Gu & Lee 2007b), Shanghai dialect (Chen 2009), and Hong Kong Cantonese (Wu & Xu 2010). Interestingly, even dialects of Chinese that do not employ F_0 manipulation for focus marking show lengthening of the syllables that are under focus.

Studies like Xu 1999 and Chen & Gussenhoven 2008 have shown for Standard Chinese that there is a stable lengthening effect within the focused word, whereas the durations hardly show any difference between neutral (no-focus), pre-focus, and post-focus condition. In other words, lengthening seems to be more local to the focused constituent, compared to the more global effect of F_0 manipulation. Furthermore, while the shape and height of F_0 contours is

influenced by a number of factors outside of focus (such as sentence position/downtrend within the sentence and the shape/height of surrounding tones), the durational influence of focus appears to be stable across different sentence positions of the focused word.

4.1.2 Focus effects and prosodic domains in Chinese

Most of the results described in the previous two sections are derived from studies that investigate focus on the word level, usually with the domain of focus being either monosyllabic or disyllabic words. However, a couple of recent studies have also asked how the focus effect is applied in longer words, or in words that serve as application domain for phonological processes, such as tone sandhi.

In Chen 2006, focus effects on the durational distribution in quadrisyllabic words in Standard Chinese were tested, including focus domains smaller than the whole word (either the disyllabic foot or the syllable). If the focus domain was smaller than the word, lengthening targeted the focus domain, and to a certain extent the immediately adjacent syllables. For example, focus on just the initial syllable of the quadrisyllabic word resulted in lengthening of that syllable, but the second syllable was also lengthened through "rightward spill-over lengthening" (Chen 2006: 197). This shows that when a focus domain undercuts a morphophonological domain, the lengthening effect may exceed the immediate focus domain.

Furthermore, the results have shown that focal lengthening applies unevenly if there is an inherent duration difference present within the focus domain. For example, in Shanghai Chinese, syllables can be intrinsically short if they end in a glottal stop or a nasal coda (Chen 2008). Under focus on disyllabic words with either inherently short or inherently long first syllables, Chen (2009) found that in words with an intrinsically long initial syllable, both syllables are lengthened under focus. In contrast, in words with an intrinsically short initial syllable, the first syllable was lengthened very little, and the second syllable received a more pronounced lengthening. This suggests some form of "compensatory lengthening" (Chen 2009), and can even be interpreted as evidence for the assumption "that durational modification for focus may be computed over the whole bi-syllabic sandhi domain" (ibid).

For F_0 effects in a focus domain below the word level, the only reported results come from Cantonese (Gu & Lee 2007b). In this study, the speaker was given disyllabic nonsense words in three different focus conditions: no focus, narrow focus on the first syllable, and narrow focus on the second syllable. The

authors report that the on-focus F_0 expansion effect of focus starts slightly before the narrowly focused syllable and decreases gradually over time, but by and large, it seems that the Cantonese speaker is able to focus only one of the syllables of the word. However, since the tested words were nonsense words, and since disyllabic compounds do not receive any special prosodic marking in Cantonese, it can still be asked whether a similar result can be found for a different dialect when testing actual compound words.

4.1.3 Wenzhou Chinese

Compared to many other dialects of Chinese, the disyllabic word domain in Wenzhou has a special status, since it serves as the application domain for tone sandhi. This means that when two syllables come together in a disyllabic compound, the tones on both syllables change in a regular, but not immediately transparent way. In that sense, the tone sandhi contour on the two syllables is functionally loaded, since it signals to the listener both information about the original lexical tones on the two syllables, and the fact that the two syllables have been compounded into a disyllabic word. If the focus effect turns out to be different in Wenzhou, this would indicate a sensitivity of focus marking to prosodic domainhood.

In some traditional impressionistic accounts of focus and tone sandhi, it has been reported that focus can block tone sandhi, for example by inserting a prosodic boundary before or after the focused constituent (Selkirk & Shen 1990; Shih 1997). Such a tone sandhi-blocking effect of focus has also been hinted at for Wenzhou Chinese (Chen 2000). However, for the Wenzhou dialect as spoken by the younger speakers today, it can be observed that tone sandhi in disyllabic words also applies under focus, even if the focus domain undercuts the tone sandhi domain. The interest of the present study is therefore to investigate whether the tonal contour that results from the application of tone sandhi still reflects the location of contrastive focus phonetically.

4.1.4 Current experiment

The current experiment is intended to investigate the effect of mismatch between prosodic domains and focus domains. The prosodic domain of interest in the current experiment is the disyllabic word, which serves as the domain for phonological tone sandhi processes in the dialect of Wenzhou Chinese. It will be systematically combined with narrow focus on five different locations with respect to the disyllabic target word.

- S1 focus: Focus on the first syllable of the disyllabic target word
- S2 focus: Focus on the second syllable of the disyllabic target word
- Word focus: Focus on the whole disyllabic target word
- Pre-target focus: Focus on the word preceding the disyllabic target word
- Post-target focus: Focus on the word following the disyllabic target word

By comparing the realizations of the tone sandhi contour in the different focus conditions, it will be investigated whether Wenzhou speakers prioritize the marking of the precise focus location, as they do by shifting the pitch accent in intonation languages, or whether the preservation of the sandhi contour as a marker of compounding will take precedence.

4.2 Methods

4.2.1 Stimuli

In order to limit the investigation of the effect of focus on a specific tone combination in Wenzhou, only disyllabic target words with a rising-falling tone sandhi contour were tested. The rise-fall tone sandhi contour results from the combination of any tone with one of the two dipping tones of the language. In this combination, any citation tone on the initial syllable will become rising, and the dipping tone on the second syllable will become falling. In contrast to level tones, contour tones (such as falling and rising tones) can be expanded in their F_0 span in both directions (i.e. upwards and downwards), and are therefore well suited for the investigation of F_0 effects of focus.

Apart from tone sandhi, the Wenzhou dialect also displays the division of the lexical tones into "registers". Broadly speaking, each tonal contour of Wenzhou (level, rising, dipping, falling) can manifest itself either in a "low" (L) or in a "high" (H) register form, co-varying with the voicing properties of the onset of the tone-bearing syllable. In order to control for register effects, all four possible combinations of registers were included in the current experiment: high register on the first and on the second syllable (HH), high register on the first and low register on the second syllable (HL), low register on the first and high register on the second syllable (LH), and low register on both syllables (LL). One of the examples that was classified as (LL beforehand was realized by the speakers as (LH), which means that there were four examples in the category (LH), two in the category (LL), and three each in the categories (HH) and (HL). All examples were compounds of the structure (noun-noun).

The tested stimulus words are listed in (2) with the following information from left to right: the Chinese characters as the speakers saw them during the experiment, a broad transcription of the target words in Wenzhou dialect, the citation tones on the first and second syllable, and a translation into English. The Chao numbers describing the citation tones are based on the description of Wenzhou in You 2002.

(2)	Register	Hanzi	Citation forms		Translation
a.	HH	中国	*tɕoŋ33*	*kai313*	'China'
b.	HH	战国	*tɕi42*	*kai313*	'Warring states period'
c.	HH	教室	*kuɔ42*	*sai313*	'classroom'
d.	HL	短袜	*tɵ35*	*mu212*	'socks'
e.	HL	中学	*tɕoŋ33*	*ɦu212*	'middle school'
f.	HL	小麦	*ɕiɛ35*	*ma212*	'wheat'
g.	LH	牙刷	*ŋu31*	*sɵ313*	'toothbrush'
h.	LH	外国	*va11*	*kai313*	'foreign country'
i.	LH	蜡烛	*ɦɛi212*	*sɛi313*	'candle'
j.	LH	语法	*nø24*	*hu313*	'grammar'
k.	LL	文学	*vaŋ31*	*ɦu212*	'literature'
l.	LL	腊肉	*la212*	*ɳɒu212*	'bacon'

The target words were embedded in a carrier sentence that remained constant across the five focus conditions, as illustrated with one of the examples in (3). This carrier sentence was coupled with a question which induced contrastive focus in one of five locations within the carrier sentence, by prompting the speaker to correct a part of the question in the answer sentence. All question-answer combinations were checked by a native speaker of Wenzhou and three native speakers of other Chinese dialects prior to the experiment to ensure grammaticality and naturalness of the sentences. In the following example, the contrasted part is marked in bold in the context questions for the reader's benefit, but it was not typographically marked for the speaker during the recording in any way.

(3) **Target sentence:**

a.	不，	我说	牙刷	商店	这几个字眼.
	fu	*ŋ kuɔ*	*ŋu sɵ*	*ɕi ti*	*ki ki kai zɻ ŋa*
	no	I say	TARGET	FRAME	these words

'No, I say the words toothbrush shop.'

S1 focus:

b.	你	说	**鞋**刷	商店	啊？
	ni	*kuɔ*	***ɦa*** *sɵ*	*ɕi ti*	*a?*
	you	say	TARGET	FRAME	Q

'Are you saying **shoe** brush shop?'

S2 focus:

c.	你	说	牙**膏**	商店	啊？
	ni	*kuɔ*	***ŋu kɜ***	*ɕi ti*	*a?*
	you	say	TARGET	FRAME	Q

'Are you saying tooth**paste** shop?'

Word focus:

d.	你	说	**短袜**	商店	啊？
	ni	*kuɔ*	***tɵ mu***	*ɕi ti*	*a?*
	you	say	TARGET	FRAME	Q

'Are you saying **sock** shop?'

Pre-target focus:

e.	你	**写**	牙刷	商店	啊？
	ni	***ɕi***	*ŋu sɵ*	*ɕi ti*	*a?*
	you	write	TARGET	FRAME	Q

'Are you **writing** toothbrush shop?'

Post-target focus:

f.	你	说	牙刷	**容器**	啊？
	ni	*kuɔ*	*ŋu sɵ*	***joŋ*** *ts^hɻ*	*a?*
	you	say	TARGET	FRAME	Q

'Are you saying toothbrush **container**?'

The disyllabic target phrase itself was part of a quadrisyllabic phrasal construction, and syntactically acted as modifier for the following disyllabic compound (glossed as FRAME in (3)). In an elicitation prior to the pilot experiment, it was ensured that in a quadrisyllabic phrasal construction like the one used in the experiment, the initial disyllabic construction would have a

clearly visible rising-falling tone contour, and not be tonally reduced due to its modifier status. The complete list of examples and precursor questions can be found in the appendix at the end of the thesis (appendix 4.1).

4.2.2 Speakers

The speakers recorded for the current experiment were all between 20 and 29 years of age (mean age = 24:0). They were all born and raised in the inner-city Lucheng district of Wenzhou, and spoke the local dialect with their friends and family on a regular basis. All of them were fluent in Standard Chinese, but had no difficulty reading out aloud Chinese characters in their dialect. None reported any hearing or speech impediments. All in all, 18 speakers were recorded for the current experiment. The data from three speakers had to be excluded from the analysis because of excessive numbers of errors and hesitations in their data, leaving the data from 15 speakers for analysis (twelve female). Five of them recorded the materials once, while the remaining ten speakers were recorded twice.

4.2.3 Experimental procedure

Speakers were recorded in a sound-proofed recording studio in Wenzhou in individual sessions, and received a small payment for their participation. Each speaker was seated in front of a 13" monitor and given a Sennheiser pc130 headset. The experimenter ensured that the microphone of the headset was placed approximately 3 cm from the corner of the mouth of the subject. Via an external digitizer (UA-G1), the sound was recorded directly on the laptop (Acer Aspire 1810TZ) on which the stimuli were displayed to the subject.

The speakers were first informed about the recording procedure. They were instructed to read out phrases and sentences presented on the screen using Wenzhou dialect in a natural and clear manner. If they were unsure how to pronounce a word or phrase, they could skip to the next item, and if they felt they had made a mistake, they could go back and repeat the recording of the previous item. They were told that they could interrupt or abort the recording at any point.

The recording itself was done using a script in the computer program PRAAT (Boersma & Weenink 2001).[14] This script presented the stimuli one by one, and recorded each stimulus individually after the speaker initiated the

[14] The script used for presenting and recording the stimuli was written by Jos Pacilly, and slightly modified by the author.

recording. Before the actual recording, all speakers completed a practice series with eight short phrases that were not part of the actual experiment. This was done in order to familiarize the speakers with the self-managed recording procedure. After completing the practice items, the speakers were asked to indicate whether they understood the recording procedure and were ready to start the actual experiment.

4.2.4 Data analysis

After the recording, all utterances were checked for mistakes and hesitations. Any pause longer than 100 ms within the carrier sentence was counted as a hesitation, and the relevant recordings were excluded from further analysis. This resulted in a total of 1424 recorded sentences which entered the analysis. All sentences were manually segmented by marking the beginning and end of the sentence, target words, and the rhymes of the target syllables, using acoustic and visual inspection of the sound wave and spectrogram. A PRAAT script measured and extracted the duration of each syllable of the target word, and performed F_0 measurements at 20 equidistant intervals within the rhymes of the target word syllables.[15] Before F_0 extraction, the measurements were checked for octave jumps and tracking anomalies due to creaky voice, and manually corrected (one octave up or down) where necessary (69 cases out of 1424 tokens).

4.3 Results

In order to assess the influence of the different focus conditions on the realization of the tone sandhi contour in the target word, several Repeated Measures (RM) ANOVAs were performed with subjects ($F1$) and items ($F2$) as random factors. Post-hoc pairwise comparisons were performed to investigate significant differences between the focus conditions, using the Sidak adjustment for multiple comparisons. The significance level adopted for the post-hoc comparisons was $p < 0.01$. All reported degrees of freedom have been Huyhn-Feldt corrected when the requirement of sphericity was not met.

[15] The script used for segmenting and measuring the files was written by Jos Pacilly.

4.3.1 F_0 effects

4.3.1.1 Register effects on F_0

In order to statistically determine whether the register on either syllable (S1, S2) interacted with the focus effects, the F_0 range on both of the two target syllables was calculated using the formula 12*ln(Fmax/Fmin)/ln(2). A by-subjects RM ANOVA on the so-calculated range as dependent variable was performed, testing the two dependent variables FOCUS (five levels) and REGISTER (four levels). This resulted in a significant main effect for FOCUS [S1 range: $F(1.28,14) = 24.74$, $p < 0.001$, S2 range: $F(2.16,14) = 38.56$, $p < 0.001$] and for REGISTER [S1 range: $F(1.52,14) = 22.74$, $p < 0.001$, S2 range: $F(2.25,14) = 32.08$, $p < 0.001$], but the two variables did not interact [FOCUS * REGISTER S1 range: $F(7.8,14) = 1.18$, $p = 0.32$, S2 range: $F(10.45,14) = 1.53$, $p = 0.13$].

The statistical test therefore shows that, while the register differences induced a significant difference in the realization of the rise-fall tone sandhi contour on both syllables, this register effect remained intact under the different focus conditions, and was not maximized or minimized in the presence or absence of focus. For this reason, Figure 4.1 displays the four register combinations on the two syllables pooled over the five focus conditions. The separate effect of focus will be discussed in the next subsection.

As can be seen in Figure 4.1, the effect of register is clearest in the early portions of the syllables, which are higher for syllables with a high register tone than for syllables with a low register tone on both syllables. As explained in section 4.2.1, this effect is a by-product of a co-occurrence constraint, according to which high register tones are realized on syllables with a voiceless onset, and low register tones are realized on syllables with a voiced onset. Consequently, on syllables with a voiced onset, the rise-fall sandhi contour will be realized lower than on syllables with a voiceless onset. The register effect is most visible in the earlier portions of both syllables, whereas it subsides in later portions of the tonal contour.

As shown in the statistical test, the register effect does not interact with the focus conditions. For this reason, the next section will discuss the focus effects averaged over all four register combinations.

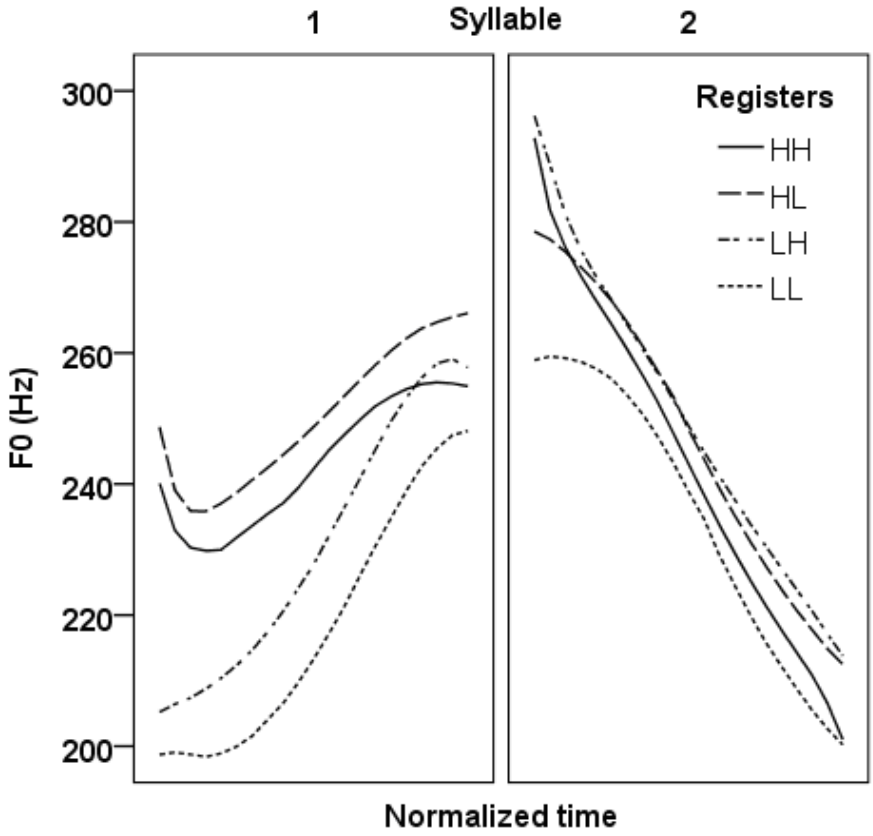

Figure 4.1: Averaged and time-normalized F_0 trajectories, pooled across focus conditions, broken down by syllable (left = syllable 1, right = syllable 2) and register combination.

4.3.1.2 Focus effects on F_0

Pooling over the different register combinations in the stimuli makes it possible to perform both by-subjects (*F*1) and by-items (*F*2) RM ANOVAs for the two syllables separately for the factor FOCUS, with F_0 range as the dependent variable. In both the by-items and the by-subjects analysis, syllable 1 and syllable 2 range showed a main effect of FOCUS [S1 range: $F1(1.22,14) = 25.65$, $p < 0.001$, $F2(2.27,11) = 89.22$, $p < 0.001$, S2 range: $F1(1.93,14) = 39.17$, $p < 0.001$, $F2(2.87,11) = 59.39$, $p < 0.001$].

Pairwise post-hoc tests across the five focus conditions showed that there was a significant difference for both syllables between the three on-target focus conditions (S1, S2, Word) and the two non-target focus conditions (Pre-target, Post-target), in that the F_0 range was expanded in the on-target focus conditions. For neither syllable did the difference between the three on-target focus conditions reach significance: S1, S2 range: S1 = S2 = Word. For two non-target focus conditions, for S1 range the by-items analysis returned a significant difference between pre- and post-target focus, but this difference failed to reach significance in the by-subjects analysis. There was no statistically significant difference between pre- and post-target focus F_0 range on the second syllable in either analysis. The findings are summarized in Table 4.1 and Figure 4.2.

Table 4.1: Summary of the statistical findings for F_0 range by focus condition. Focus conditions: S1= Syllable 1 focus, S2 = Syllable 2 focus, Word = Word focus, Pre = Pre-target focus, Post = Post-target focus.

	Syllable 1 range	**Syllable 2 range**
By-subjects (*F*1)	S1, S2, Word > Pre, Post	S1, S2, Word > Pre, Post
By-items (*F*2)	S1, S2, Word > Post > Pre	S1, S2, Word > Pre, Post

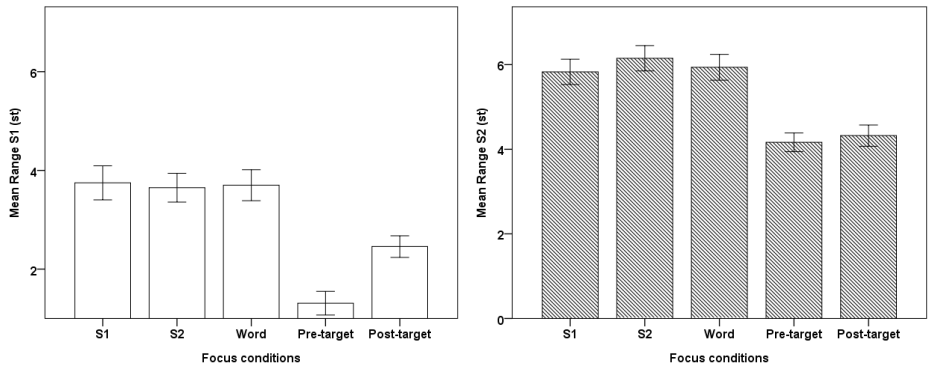

Figure 4.2: Mean absolute pitch change and error bars (±2 SE) for the first (left) and second (right) syllable across five focus conditions. Focus conditions (from left to right): S1 focus, S2 focus, Word focus, Pre-target focus, Post-target focus.

In order to investigate why pre- and post-target focus yield a difference in F_0 range on the first, but not on the second syllable, it is beneficial to look at the actual F_0 curves of the five focus conditions. As can be seen in Figure 4.3, the rising tone on the first syllable in pre-target focus condition starts higher than in all other conditions, and falls a bit before starting to rise later than in the other conditions, thereby reaching a lower F_0 maximum. This is a confound from the fact that in pre-target focus, the syllable preceding the target word, which bears a rising tone, is focused.

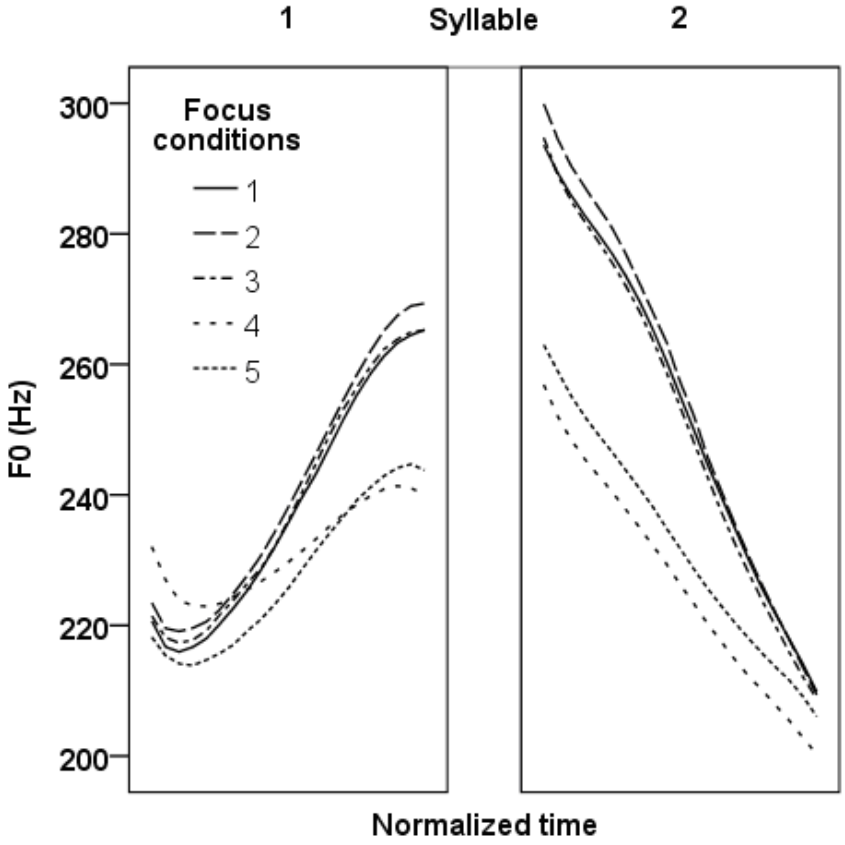

Figure 4.3: Averaged and time-normalized F_0 trajectories across five focus conditions, broken down by syllable (left = syllable 1, right = syllable 2) and focus conditions (1 = S1 focus, 2 = S2 focus, 3 = Word focus, 4 = Pre-target focus, 5 = Post-target focus).

Under focus, this rising tone is realized with a greater F_0 range than in the other four conditions, which is why the tonal contour on the following syllable has to fall first before rising again. After the coarticulatory effect of the previous syllable has worn off, the tonal contour for the pre-target focus is quite similar to that of post-target focus on the second syllable. It can be assumed that in absence of this confound, the pre- and post-target focus condition would also yield comparable effects on the first syllable.

In Figure 4.3, it can furthermore be seen that he F_0 expansion effect of focus seems to be unidirectional. The F_0 range on the focused constituents is expanded mostly upwards, whereas the F_0 minima remain comparable across the five focus conditions. To test whether this impression holds statistically, the F_0 maxima and minima were also compared across the different focus conditions for both syllables. The results are shown in Figure 4.4 below.

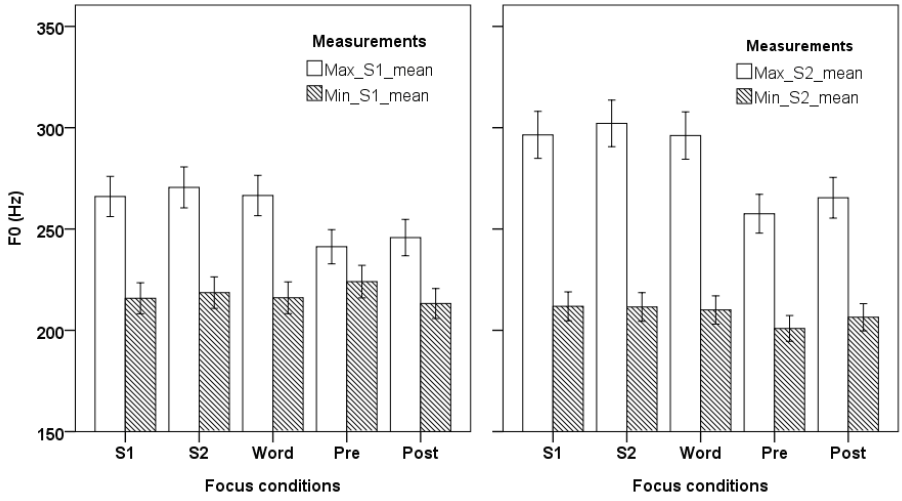

Figure 4.4: Mean F_0 maxima (clear boxes) and minima (patterned boxes) for the first (left) and second (right) syllable across five focus conditions. Focus conditions (from left to right): S1 focus, S2 focus, Word focus, Pre-target focus, Post-target focus. T-bars = ±2 SE.

Two RM ANOVAS, again one by subjects (*F*1) and one by items (*F*2), both yielded main effects for the factor FOCUS on both syllables for all measurements: Syllable 1 F_0 maximum [$F1(1.56,14) = 35.63$, $p < 0.001$, $F2(3.66,11) = 100.03$, $p < 0.001$], Syllable 1 F_0 minimum [$F1(1.41,14) = 4.65$, $p = 0.032$, $F2(2.14,11)$

= 7.69, p < 0.003], Syllable 2 F_0 maximum [*F*1(1.43,14) = 48.07, p < 0.001, *F*2(3.91,11) = 140.75, p < 0.001], and Syllable 2 F_0 minimum [*F*1(2.98,14) = 9.95, p < 0.001, *F*2(2.41,11) = 12.66, p < 0.001].

Post-hoc pairwise comparisons across the five focus conditions, as illustrated in Table 4.2, revealed a similar picture for the F_0 maxima on both syllables as the F_0 range measurements: the F_0 maxima were significantly different between the three on-target focus conditions and the two non-target focus conditions, but there was no significant difference within either group for either analysis. For the F_0 minima, there was no such clear division between the different target conditions, and only some spurious significances surfaced, mostly involving the raised F_0 minimum on the first syllable of the target word due to the preceding rising tone.

Table 4.2: Summary of the statistical findings for F_0 range by focus condition. Focus conditions: S1= Syllable 1 focus, S2 = Syllable 2 focus, Word = Word focus, Pre = Pre-target focus, Post = Post-target focus.

	By-subjects (*F*1)	By-items (*F*2)
S1 F_0 maximum	S1, S2, Word > Pre, Post	S1, S2, Word > Pre, Post
S1 F_0 minimum	S2 > Post	Pre > Post S2 > Post
S2 F_0 maximum	S1, S2, Word > Pre, Post	S1, S2, Word > Pre, Post
S2 F_0 minimum	S1, S2, Word > Pre	S1, S2, Word, Post> Pre

The results show that the significant effect of focus on the F0 range on the two syllables of the target word is brought about more or less exclusively by a raising of the F0 maxima, whereas the F0 minima remain relatively unaffected by the focus conditions, and only show secondary coarticulatory effects. For the F0 minima, there is no systematic way to distinguish all on-target focus conditions from all non-target focus conditions, which shows that the F0 expansion effect of focus in Wenzhou is unidirectional.

4.3.2 Duration effects

For the duration measurements of both syllables, again a by-subjects (*F*1) and by-items (*F*2) RM ANOVA was conducted with FOCUS as dependent variable. Both analyses returned main effects of FOCUS on the duration measurements: Syllable 1 duration [$F1(3.27,14) = 25.49$, $p < 0.001$, $F2(4,11) = 15.75$, $p < 0.001$], and Syllable 2 duration [$F1(2.25,14) = 36.17$, $p < 0.001$, $F2(3.72,11) = 41.2$, $p < 0.001$]. The measurements are illustrated in Figure 4.5 below.

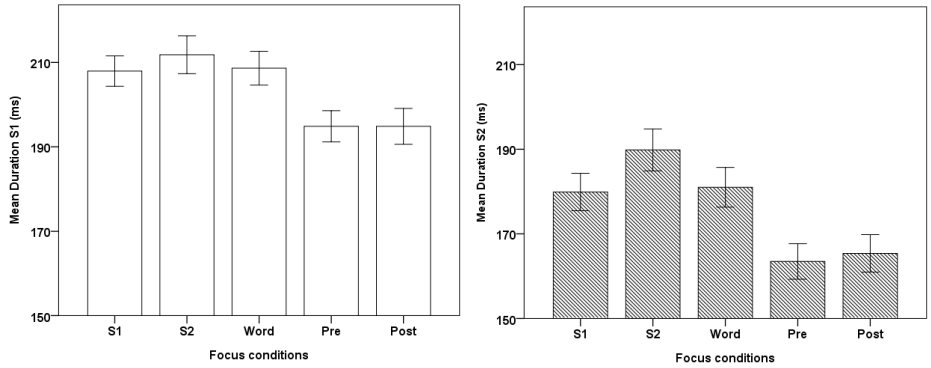

Figure 4.5: Mean duration of syllable 1 (left) and syllable 2 (right) of the target word across five focus conditions. Focus conditions (from left to right): S1 focus, S2 focus, Word focus, Pre-target focus, Post-target focus. T-bars = ±2 SE.

The post-hoc pairwise comparisons across the five FOCUS conditions again showed a significant difference between the three on-target focus conditions (S1, S2, Word) and the two non-target focus conditions (Pre-target, Post-target) on the first syllable, in that its duration was longer under focus. However, the detailed comparisons on the second syllable returned different results for the two analyses. While in the by-items analysis there was no significant difference between the three on-target focus conditions in either measurement (S1 = S2 = Word), the by-subjects analysis showed a significant difference for the length of the second syllable: it was longer under narrow focus (S2 focus) than when focus was on the first syllable (S1 focus).

However, the difference in duration of the second syllable under S2 focus did not reach significance in comparison with the Word focus condition. Pre- and Post-target focus did not return a significant difference in either analysis. The findings are summarized in Table 4.3.

Table 4.3: Summary of the statistical findings for syllable duration by focus condition. Focus conditions: S1= Syllable 1 focus, S2 = Syllable 2 focus, Word = Word focus, Pre = Pre-target focus, Post = Post-target focus.

	Syllable 1 duration	**Syllable 2 duration**
By-subjects (*F*1)	S1, S2, Word > Pre, Post	S2 > S1 S1, S2, Word > Pre, Post
By-items (*F*2)	S1, S2, Word > Pre, Post	S1, S2, Word > Pre, Post

4.4 Discussion

4.4.1 Summary of results

The results of the experiment show that in Wenzhou Chinese as in other Chinese dialects, contrastive focus is marked both by F_0 range expansion and lengthening. However, neither of the two effects allows for a systematic distinction between the three on-target focus conditions, or between the two non-target focus conditions.

From an F_0 perspective, the phonetic effects of focus on the first syllable, the second syllable, or the entire target word within the disyllabic tone sandhi domain of Wenzhou Chinese are similar. This means that, in Wenzhou Chinese unlike in many intonational languages, focus cannot pick out one syllable in a word as the location of focus and mark it with a distinct F_0 movement to distinguish it from other syllables within that word. Rather, the F_0 contour that results from disyllabic tone sandhi remains intact under focus: even when focus undercuts the tone sandhi domain, the phonetic reflex of focus is similar to when the entire word is in focus.

In terms of duration, there is a clear difference between the three on-target focus conditions and the two non-target focus conditions, which manifests itself in a stable lengthening effect on both syllables. Additionally, under S2 focus, the second syllable is lengthened to a greater extent than under S1 focus, while no comparable lengthening effect of S1 focus is visible on the first syllable. However, the duration of the second syllable under S2 focus is not significantly different from the duration of the same syllable under Word focus, and only reaches significance compared to S1 focus in one of the two statistical

tests. This leads to the conclusion that, while there is slightly greater lengthening on the second syllable when it is in narrow focus, the duration effects on neither syllable are big enough to allow distinction between all three on-target focus conditions.

Besides the absolute lengthening effect of focus on both syllables, it is also interesting to look at the relative duration of the two syllables with respect to one another. As can be seen in Figure 4.5, it is the case in all the focus conditions that the duration of the first syllable surpasses that of the second syllable. This is comparable to the edge-effect described in Chen 2006, which states that word-initial syllables tend to be longer than medial syllables (*ceteris paribus*). In the current experiment, the quadrisyllabic phrase containing the disyllabic target word and the disyllabic noun which it modifies, shows a comparable pattern of internal length distribution to the quadrisyllabic words in Standard Chinese described in Chen 2006.

However, the effect of lengthening under focus is different from that described for Standard Chinese, in that the durational distribution within the target word (initial syllable always longer than second syllable) remains intact under all focus conditions. In contrast, in Chen 2006, the duration of the second syllable under S2 focus exceeded that of the first syllable. This finding presents further evidence for the conclusion that the lengthening effect of focus in Wenzhou is distributed over both syllables of a disyllabic word, regardless of the exact position of the focus within that word.

As for the pre-and post-target conditions, they prove indistinguishable, except where external factors affect either part of the measurement domain (cf. Figure 4.3). When looking at the syllable 1 F_0 range, it is smaller in pre-target focus condition than in post-target focus condition, because the (focused) rising tone on the pre-target syllable expands its F_0 range and thereby diminishes the F_0 range of the immediately following syllable. However, when comparing the F_0 maxima and minima independently, it can be seen that only the F_0 minima are affected by this coarticulation, whereas the F_0 maxima only reflect focus condition. Furthermore, the present experiment finds neither a lengthening nor an F_0 effect that distinguishes material in pre- from post-focus position.

4.4.2 Implications of the findings

The current study has shown that the phonetic mechanisms of marking focus on sandhi tones, namely F_0 range and duration expansion, are similar to those that have been reported for lexical tones, a finding that was already discussed for the sandhi tones in Shanghai Chinese in Chen 2009. However, unlike in Shanghai

Chinese, no significant post-focal lowering of F_0 contours was found in the present experiment. This gives reason to conclude that Wenzhou, unlike Standard Chinese and Shanghai Chinese, does not employ post-focal lowering as an additional cue to focus location.

In Wenzhou Chinese, the realization of contrastive focus in the disyllabic word domain, which is the domain of phonological tone change in this language, is influenced by the specific characteristics of this domain. What has been shown is that the focus effects of F_0 range expansion and lengthening are distributed over the entire disyllabic domain, even when only one of the two syllables in this domain is the target of the contrastive focus. This is in contrast to findings for other dialects of Chinese, such as Standard Chinese (Chen 2006) and Hong Kong Cantonese (Gu & Lee 2007b), for which it has been shown that contrastively focused syllables within polysyllabic words can receive greater phonetic marking than other syllables in the same word.

By contrast, in Wenzhou Chinese, it appears that the disyllabic tone sandhi domain cannot be split up by contrastive focus, so that one syllable would receive greater phonetic marking than the other. Rather, both the F_0 contour and the duration distribution of the disyllabic words are similar under focus on the whole word or on one of its parts. It appears that focus can only affect the entire disyllabic tone sandhi domain as a whole, but not break up its internal structure.

This is in contrast to what has been found for intonation languages such as Dutch, where the marking of the exact location of focus can take precedence over the phonetic display of phonological properties such as lexical stress, and for example shift the location of a pitch accent within a word. Via the segmental durations, the speaker still receives cues to locate the lexical stress properly, but the primary cue, namely the pitch accent, is utilized for focus marking. In Wenzhou Chinese on the other hand, it seems that the realization of the tone sandhi F_0 contour mainly serves to convey the cue of disyllabic wordhood (rather than just two syllables that happen to be adjacent), at the expense of marking the exact location of contrastive focus for the listener.

While there is a minimal effect of duration, which is extended on the second syllable when this syllable is narrowly focused, it seems that also the durational marking is mostly uniformly expanded over the entire disyllabic domain under focus. This is similar to what has been found for Dutch (Sluijter & van Heuven 1995), namely that while focus induces a lengthening of the word under focus, the internal durational distribution between the syllables of the word remains mostly intact. In that sense, speakers receive a (durational) clue

for stress in Dutch even under focus, and analogously it can be said that they receive a (durational) clue for the phrase-initial position (i.e. the first syllable being longer than the second), even when focus marking would favor lengthening on the second syllable.

These findings can be interpreted to mean that the speakers conceptualize the entire disyllabic tone sandhi domain as one whole, and that the integrity of the domain is preserved under focus. Rather than individually expanding the F_0 range or duration of one of the syllables, the speakers expand the contour on both syllables upwards to strengthen the tonal realization. This finding also speaks for an interpretation of the tone sandhi process whereby, as soon as two lexical tones come together in a disyllabic compound word, these tones are "replaced" by a tonal contour that is spread over the whole disyllabic word domain. Additional effects such as focus marking can then only affect this contour as a whole, but not break it up into its components any more to emphasize one over the other.

At the same time, the findings speak for a view of focus marking that has to allow a greater room for phonological processes or prosodic constituents. While it seems that the straightforward effects of focus on F_0 and duration lend themselves to an analysis that sees focus as something that is phonetically implemented, such a conclusion would be at odds with the results of the current experiment. If focus were just a phonetic effect that gets added to the finished derivation, it would be counterintuitive to expect that it should pay attention to the lexical integrity of compound words in one dialect (Wenzhou), but not in others (Standard Chinese).

Rather, it seems that the current findings lend themselves to a more indirect view of focus (see also Chen 2009 and Chen & Gussenhoven 2008), which sees the effects of focus as comparable to the strengthening effect of prosodic prominence within prosodic constituents. Under such a view, the F_0 range and duration expansion observed under focus is then a consequence of a more abstract, phonological "strengthening" effect brought about by focus, rather than the manifestation of a phonetic focus effect itself. Such a more indirect effect of focus, which is mediated by prosodic structure specifications in the respective language, can also help to explain recent findings for multiple focus in Standard Chinese (Kabagema-Bilan et al. 2011).

For the present experiment, it is clear that prosodic structure plays an important role in the implementation of focus, namely by constraining how narrowly focus can be marked. The results here suggest that the tone sandhi domain, within which changes to the F_0 contour of syllables within disyllabic

words take place, is also the domain that limits the distribution of focus marking, at least when it comes to F_0 effects. When computing the tonal contours of disyllabic compounds, Wenzhou speakers have to take the tonal information on both syllables into account, and see the entire disyllabic word as one whole. This holistic perspective is reflected in focus marking by F_0 expansion, which applies uniformly across the two syllables within the domain, no matter whether the actual focus domain is the whole word or either of the two syllables.

It is interesting to compare the findings of the current experiment to the (brief) description of focus effects in Wenzhou in Chen 2000. In this book chapter, which is based on the impressionistic description of recordings from a middle-aged speaker in the 1980s, it is not only predicted that focus should be able to single out individual syllables in the disyllabic compound domain, but also that it should be able to break up the disyllabic prosodic domain at all, and interrupt the phonological process of tone sandhi within it (so that each syllable would be realized individually with its lexical tone).

Such a phonological view of focus, in which the focus can directly manipulate the presence/absence of prosodic boundaries and the phonological processes that are connected to them, has been showed to be inaccurate on other grounds before (Chen 2004). For the young speakers recorded in the current experiment, it seems that the limitations of focus marking are yet one step further ahead: not only does the tone sandhi contour remain intact in the presence of a focus that singles out an individual syllable within it, but even the phonetic implementation of the focus effect appears to be mediated by the prosodic tone sandhi domain as a whole.

Therefore, the need for a more "phonological" view of focus effects in Chinese should not be taken to mean that focus should be able to modify the prosodic structure as it is mapped from syntax. Rather, focus appears to be sensitive to the prosodic structure and its specifications, but unable to change its direct components. Rather, this prosodic structure seems to limit the extent to which focus can modify the tonal information, while still ensuring that crucial parts of the information (such as "wordhood" in the current experiment) remain intact.

4.5 Conclusion

This study investigated the distribution of narrow focus marking within the disyllabic tone sandhi domain of Wenzhou Chinese. An experiment looked at the influence of sub-word focus on the implementation of the rise-fall sandhi

contour in disyllabic words. Recordings from 15 young Wenzhou speakers were analyzed, in which they read out question-answer pairs which induced contrastive focus on either or both syllables of the disyllabic target word. Additionally, focus on the pre- and post-target word was tested for comparison.

Analysis of the F0 curves of the recordings shows that, regardless of the exact location of focus within the disyllabic tone sandhi domain, the tonal contours on both syllables are modified in a similar manner, compared to the control conditions (i.e. pre- and post-target focus). Lengthening likewise targets both syllables under focus on the whole word, and there is no significant difference between word focus and focus on either syllable. The second syllable shows slightly greater lengthening under syllable 2 focus than under syllable 1 focus, but this duration difference is not enough to reliably distinguish all three focus conditions from one another.

These findings suggest that focus affects the realization of the entire tone sandhi contour, even when only one of the syllables is contrasted. Sandhi tones are derived within the disyllabic domain, and likewise the F0 effects of focus are spread out over the entire domain. For duration, the only difference between the conditions is on the second syllable when it is narrowly focused, but the small size of the effect suggests that the entire disyllabic tone sandhi domain can still be considered as the location of the prosodic implementation of focus effects. In short, the disyllabic tone sandhi domain limits the distribution of focus effects in Wenzhou Chinese, which argues for a non-direct mapping of focus domain with the prosodic marking of focus.