Based on the framework developed in Halle & Vergnaud (1987), a two-way model of metrical analysis is proposed to resolve the prosodic tension between sandhi and stress as documented in Wenzhou Chinese. In particular, while following a last-dominant paradigm in its disyllabic sandhi, the language concerned chooses to go awry by letting stress descending on the initial syllable which, as the target of sandhi, is phonologically weak. The resultant trochaic contour of word stress is thus outrageously in conflict with the last-dominant sandhi pattern. This gives rise to what we call the prosodic tension here. The present model offers a resolution of this tension by proposing a metrical analysis as follows. That is, the metrical derivation of a disyllabic word begins with an iambic parsing algorithm, which conditions the last-dominant paradigm of sandhi in the language. In the process of derivation, however, the initial tonal upheaval effected by a high falling tone sandhi brings in its wake vocalic lengthening, a process of iambic lengthening, resulting in the deformation of disyllabic metrical structure. Under the combined force of foot binarity and quantity affinity inherent in iambic prosody, the iambic lengthening leads to refooting of the deformed metrical constituent. Remetrification then follows the exactly same iambic parameter settings as the original parse, but is critically preceded by an iambic accenting, which is to guarantee that the derived heavy syllable constitutes an independent foot on its own.

1. Introduction

This paper is a metrical analysis of disyllabic tone sandhi in Wenzhou Chinese. The language is spoken mainly in Wenzhou, a city with a total population of about 1 million, 400 miles south of Shanghai. Genetically it is a variety of the Wu dialect, a linguistic family representing a geographic area extending from the southern part of Jiangsu Province to most of Zhejiang Province. Though the city is geographically somewhat isolated, the language spoken there has been said to be an "impeccable"
representative of the Wu group for the reason that it possesses all those classifying properties that marks off all and only Wu dialects among a myriad of Chinese languages and dialects. Chao (1967) extracted three defining phonological features from among the Wu dialects, two of which are relevant to the present discussion. First, the Wu dialects tend to raise the low and mid vowels of cognate words in non-Wu dialects, often accompanied by a simplification of diphthongs. Second, the Wu dialects have a richer inventory of lexical tones than any other language groups in the nation.

Wenzhou brings these two classificatory features into full play. First, it has the richest inventory of lexical tones. In particular, while most of other Wu dialects have 5 to 7 tones Wenzhou has eight lexical tones by splitting the ancient 4 grand tonal categories (i.e. Ping, Shang, Qu, Ru,) into two parallel series, namely Yin versus Yan. The contrast of Yin (literally female) versus Yan (literally male) is originally a philosophical notion in Sinology. Chinese phonologists borrow the notion to capture register contrast of tone by assuming that Yin is phonologically high while Yan low. Second, vowels in Wenzhou doe not contrast in length by reducing diphthongs to pure vowels. More precisely, the vowels are all underlyingly short.

With its striking phonological characteristics, Wenzhou has recently inspired quite a few North America-based studies (Bao, 1990, 1997; Chen, 1990, 1996; Duanmu, 1990; Yip, 1995 to name just a few). All these studies, at some point, and, of course, to a different degree, show their concern with the issue of tone sandhi and stress. But none of them seems to have been genuinely impressed by one particularly striking phenomenon in the language that is conspicuously documented in the same source which all those authors have extensively drawn upon. Specifically, the source is Zheng-zhang (1964a & b), who notes, in several places, that there is a mismatch between tone sandhi and stress assignment. In particular, the overwhelming majority of Wenzhou disyllabic combinations follow a last-dominant sandhi paradigm, meaning that the last syllable tends to preserve its citation tonal pattern while the initial one undergoes drastic tonal change. Given the fact that the final syllable is phonologically strong and thereby conditions the initial sandhi, the word stress contour over a
Resolving Tension between Tone and Stress

disyllabic should be expected to demonstrate an iambic rhythm. However, "Contrary to our expectation", as noted by the author,

it is not the case. Instead it is the first syllable that is perceived as somewhat stronger than the second one. Thus when one wants to emphasize the disyllabic, this can be achieved only by accenting the first syllable, not the second one, though the former has completely lost its lexical tone (Zheng-zhang, 1964b: 109).

As a native speaker of the language, I share the author's perceptual judgement. To all appearances, the overwhelming majority (to be made precise in the body of the paper) of disyllabics in Wenzhou represent an outright contradiction to a basic linguistic assumption, that is, it is typically the phonologically strong, not the weak, position that preserves structure. Since stressed position is believed to be phonologically strong, it should be a stressless syllable that will be subject to tone sandhi. Surprisingly, however, most of Wenzhou disyllabics just choose to go awry by letting tone sandhi rules target a stressed syllable.

For ease of reference, I describe this situation as prosodic tension. This paper proposes a resolution of the tension by advancing a two-way model, whereby tone sandhi rules and stress assignment interact in a manner such that the output from one element serves as input to the other. In particular, metrical footing precedes the application of tone sandhi rules, and then the effects resulting from tone sandhi serve as input to metrical refooting under force of some universal metrical principles, the output of which in turn constitutes input to tone neutralization on the final syllable. Summarily, I contend that it is a refooting twist that gives rise to a distorted surface stress contour over Wenzhou disyllabic words, that is, a final-dominant sandhi pattern but with a trochaic stress contour.

The remainder of the paper is organized as follows. Section II defines the problem, proceeding from a rough sketch of disyllabic tone sandhi across Chinese dialects to a precise description of Wenzhou disyllabic sandhi patterns. Section III spells out the details of metrical derivation, explaining how tone sandhi comes to interact with metrification. Section IV concludes the paper with a summary.
2. Defining the Problem
Zheng-zhang's work was done in early 60s and essentially confined to China. His presentation of prosodic tension between tone and stress indicates that he has apparently taken stress as a segmental feature. He simply had no access to the basic insight of more recent metrical theory: stress is not a feature, but an abstract structure wherein rhythmical positions are organized into headed constituents. A straightforward extension of this fundamental tenet is then that metrical structure is not necessarily realized through phonetic contrast. This statement is particularly meaningful when one is dealing with Chinese languages wherein phonetic word stress is rarely contrastive lexically. This has made linguists working on Chinese phonology seek metrical evidence elsewhere. This elsewhere, to my knowledge, is almost exclusively on the relation between stress and tone sandhi.

Starting with Yip (1980), the studies of the relationship between tone and stress have followed one of two opposing positions: either stress conditions tone sandhi or vice versa. Regardless of the technical difference between them, however, the two opposing positions share the same methodological ground. That is, tone and stress interact with each other in a one-way mode such that "n any case, there is a clear connection between the inability of toneless syllables to bear stress..., and the inability of stressless syllables to bear tones ...." Yip (1995: 490).

While faring well with some data from some Chinese dialects, this one-way model is unable to accommodate Wenzhou data. To better define the problem, I digress to a brief sketch of disyllabic sandhi across Chinese dialects, to provide a minimal background for an appropriate characterization of Wenzhou disyllabic sandhi.

Chinese is by and large a monosyllabic language--monosyllabic in the sense that each syllable can be isolated as a morpheme, which carries a distinct tone. Thus, it is theoretically possible to identify each single syllable with a tone, which is termed ban diao (literally, original tone) in the literature of Chinese phonology, or the citation / lexical tone in our terminology. When syllables are strung together the citation tone of some of the syllables in the string will undergo a change or modification, which is termed bian diao (literally, change tone) by Chinese linguists. So it is
quite straightforward to define tone sandhi as systematic difference in the tone value of a syllable in a string vis-à-vis the tone value of the same syllable uttered in isolated citation form. Reasoning along this line, it is equally quite clear that a disyllabic string is minimally required for a tone sandhi to happen. As a side note, this paper regards disyllabic strings as a "purely" phonologically defined sandhi domain leaving out the much-discussed issue of mapping between minimal sandhi domain and morphological/syntax domain.

Typologically speaking, majority Chinese dialects just simply follow one of the two types of tone sandhi paradigms: first-syllable versus last-syllable dominance (Hashimoto, 1987). In the former type, the first syllable has the same or similar tone value as that of its citation form while the second one assumes a different tone, or sandhi tone. In the last-syllable dominant type, the reverse is true, namely the sandhi target is on the initial syllable while the last syllable preserves its citation form (Hashimoto, 1987). Following current derivational framework, these two types of tone sandhi paradigms can be schematically represented as follows:

(1) Initial dominance: $\sigma^T + \sigma^T \rightarrow \sigma^T + \sigma \rightarrow \sigma^T \sigma^S$

Last dominance: $\sigma^T + \sigma^T \rightarrow \sigma + \sigma^T \rightarrow \sigma^S \sigma^T$

$T =$ citation tone;
$S =$ sandhi tone

Embedded in the representation above is one fundamental question. That is, what is the precise process that effects tone sandhi? Without committing myself to a fine-grained picture, my answer to the question is this: sandhi tone is effected by either pattern extension or pattern substitution. The terms are due to Chan & Ren (1986), and adopted here mainly for their theory-neutral connotation. As far as the goals of the present paper are concerned, I simply cannot afford to commit myself to

---

2 Tone change does happen on one single syllable, which usually carries some morphological meaning (e.g. diminutive). Chinese linguists distinguish this kind of tone change from tone sandhi (i.e. bian diao) by dubbing the former bian yin (literally, change sound). I follow this practice by seeing disyllabic as the minimal domain of tone sandhi, to reflect one fundamental fact: tone in Chinese is almost exclusively used lexically, with little morphological function, as compared to extensive morphological load that tone carries in many other tone languages in Africa and America.
one or other sandhi theory without losing my focus.

By pattern extension is meant a process of tonal spreading from the dominant syllable to the dominated syllable that has undergone a deletion of its original tone. Thus, the end result of pattern extension is such that the pitch contour of a disyllabic combination as a whole is entirely dependent on the citation contour of the dominant syllable regardless of whatever the inherent tone the dominated syllable originally assumes. For example, suppose that the initial syllable is dominant with a high falling tone while the second syllable carries one through four distinct citation tones respectively, then the surface pitch contour of all these four words will be identical, starting high and finishing low. This is exactly what has been extensively attested with disyllabic, and in fact, all polysyllabic tonal pattern in Shanghai Chinese (a Wu dialect). This observation should be credited to Sherard (1972). But I quote Yip (1980:59) for illustration of pattern extension:

(2) pattern extension of tone sandhi

\[
\text{thy}^{51} + \text{sang}^{51} \quad \rightarrow \quad \text{thy sang} \\
\text{kong}^{51} + \text{si}^{44} \quad \rightarrow \quad \text{kong si} \\
\text{thy}^{51} + \text{seq}^{213} \quad \rightarrow \quad \text{they seq} \\
\text{sang}^{51} + \text{wheq}^{5} \quad \rightarrow \quad \text{sang wheq}
\]

"naturally"

"company"

"weather"

"life"

(Yip represents tones using Chao's digital system, where 1 is lowest in pitch and 5 is highest. This tonal representation system will be followed in this paper.)

Pattern substitution is more elusive. Like pattern extension, the dominant syllable preserves its citation form, but the non-dominant syllable picks up a tone that is independent of both the original form of its own, and the
Resolving Tension between Tone and Stress

citation form of the dominant syllable. Intuitively, we might as well, happily or unhappily, call this sandhi form substitute tone. Hence the term pattern substitution. While allowing for various theorizing, the label substitute tone is unappealing within generative phonology, as pointed out by Yip (1995). However, at the same time, Yip herself has no solution for the problem. As a matter of fact, while coming to typology of sandhi in Chinese, she becomes rather fuzzy by saying "there is a fundamental division between tone changes caused by a specific tonal environment ... and tone changes caused by purely positional factors..." (1995: 491). One never knows what is precisely meant by "purely positional." My interpretation is that pattern substitution should be one instance of the so-called purely positional phenomenon, since it usually occurs on domain edge, as will become clear in what follows.

As a first illustration of pattern substitution in sandhi, I turn Tone 3 sandhi in standard Mandarin Chinese. In particular, standard Mandarin is believed to have four lexical tones (i.e. Tone 1 through Tone 4). The Tone 3 sandhi rule can be represented as follows:

(3) Tone 3 sandhi in standard Mandarin

Tone 3 --- > Tone 2 / _____ Tone 3
(Ex. hao3 + jiu3 --- > hao2 jiu3)

good wine good wine

What the sandhi rule says is that Tone 3 on the initial syllable is replaced by Tone 2 when two syllables with identical Tone 3 are juxtaposed to each other. The assumption here is transparent enough: the second syllable is dominant and thus conditions a pattern substitution on the initial one. Please note that the sandhi tone (i.e. Tone 2) as the output of rule (3), is an inventory lexical tone in the language. This amounts to saying that pattern substitution is a paradigmatic replacement. As indicated earlier, paradigmatic replacement is an unappealing notion, but generative theory does have a way of explaining it away in this instance. This is the principle of OCP, whereby a statement like the following is largely acceptable: Tone 3 alternates with Tone 2 when juxtaposed before an identical Tone 3. 3

---

3 The nature of contour tones is yet to be resolved: one view takes a contour tone as a unitary
However, there is data on pattern substitution that cannot be handled by the OCP. The Min dialects, for example, have been thoroughly attested for their characteristic pattern substitution (Yip, 1995). All the substitutions there are "purely positional". In particular, disyllabic combinations in this group of dialects, like the Mandarin example in (3), follow a last-dominant paradigm, but almost every lexical tone on the last syllable has a matched sandhi form on the initial syllable. Like the Min dialects, Wenzhou follows a last-dominant sandhi paradigm. But unlike the Min dialects, there is an extensive merger of initial sandhi forms in Wenzhou, resulting in only two types of disyllabic sandhi forms, which I will dub Type I and II respectively for convenience. Type I is of those disyllabics whose initial syllable invariably surfaces as a low-level short tone, which is usually transcribed as 1 in Chao's number. Type II complimentarily subsumes the rest of disyllabic combinations other than those of Type I in the language. The common thread among Type II disyllabics is that their initial syllable surfaces up as a high falling tone, which is transcribed as 42 or 53, depending on the final syllable. Taxonomically, I list these two types of initial sandhi forms in (4) below.

(4) Two types of Wenzhou disyllabic sandhi

<table>
<thead>
<tr>
<th>Type</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>T + T</td>
<td>1 + T</td>
</tr>
<tr>
<td>II</td>
<td>T + T</td>
<td>42/53 + T</td>
</tr>
</tbody>
</table>

The schema in (4) is at best a remote representation of Wenzhou disyllabic sandhi. Critically, there is one important complication in Wenzhou sandhi pattern. In particular, the last dominant syllable itself, while conditioning tone sandhi on the initial syllable, also undergoes a certain degree of tonal modification. By certain degree is meant the tonal change on the last syllable is limited to a well-defined category -- the category of tonal register contrast which is known as Yin vs. Yan in traditional Chinese phonology.

To account for this phenomenon, let me highlight one traditional insight into tonal organization in Chinese. That is, one of the organizing

melodic gestalt while the other argues that it is a compositional cluster.
Resolving Tension between Tone and Stress

principles of tonal contrasts in Chinese is traditionally couched in terms of Yin vs. Yan. This refers to a contrast in tonal register, derived from voicing contrast on the initial of the syllable, with Yin from voiceless initials and Yan from voiced initial. The term Yin vs. Yan indicates that Chinese scholars have been 17 centuries ahead of the Prague school in formalizing a phonological contrast to accommodate the infinite world of phonetics. Yip (1980) captures the bilateral nature embedded in the opposition of Yin vs. Yan by the feature Upper, a binary-valued feature that enables us to interpret Yin as [+upper] and Yan as [-Upper]. With this notion in hand, I am now ready to present the Wenzhou tonal system.

(5) Inventory of Wenzhou citation tones

<table>
<thead>
<tr>
<th>Middle Chinese Tone Categories</th>
<th>I. “Ping” Even</th>
<th>II. “Shang” Ascending</th>
<th>III. “Qu” Departing</th>
<th>IV. “Ru” Entering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wenzhou Citation Tones</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yin (+Upper)</td>
<td>Ia 33</td>
<td>IIa 45</td>
<td>IIIa 43</td>
<td>IVa 313</td>
</tr>
<tr>
<td>Yan (-Upper)</td>
<td>Ib 31</td>
<td>Iib 34</td>
<td>IIIb 11</td>
<td>IVb 212</td>
</tr>
</tbody>
</table>

Like all other Chinese dialects, Wenzhou possesses four primary tonal categories that are believed to have been taken over from Middle Chinese. These ancient tonal categories are known as Ping [even], Shang [ascending], Qu [departing], and Ru [entering]. Wenzhou’s claim to fame among Chinese dialects is that it dissects the four primary tonal categories into Yin vs. Yan contrast to perfection, resulting in a neat inventory of 8 citation tones. In the table above, the four primary tonal categories are numbered sequentially from I to IV, each of which is then split into Xa and Xb, reflecting Yin vs. Yan contrast. Thus, for example, Ila is termed Yin Shang while Iib Yan Shang. Since they belong to the same grand category (i.e. Shang), these two tones in principle share the same melodic contour (i.e. rising contour). They contrast in register only: one (i.e. Ila) is higher

---

4 The practice of splitting syllables into initial (=onset) plus final (=rhyme) dates back as early as to 600 AD in the history of Chinese phonology. There have been ample discussions on how well this ancient notion fits into modern concepts of syllables (e.g. Bao, 1990).
Shane Xuexin Cao

(thus named Yin), and the other (i.e. IIb) is lower (thus named Yan). Or in Yip's words, they contrast because one is [+ Upper] while the other is [-Upper]. Please note that register contrast is always bilateral within each grand category. To provide a reference of the pitch shape, all the resultant 8 tones are annotated in Chao's tonal digits, based on the original source (Zheng-zhang, 1964b, 1980).5

With the inventory of citation tones listed and explained, let me return to the taxonomic representations of disyllabic sandhi in (4), which I am now in a position to refine.

(6) disyllabic tone sandhi in Wenzhou

Type I: \( T^{Ru} + T \rightarrow 1 + T \)

Type II: \( T^{non-Ru} + T \rightarrow 42/53 + T \)

\( T \) = citation tone

\( T^{Ru} = 213 \) or 212

\( T^{non-Ru} \) = the remaining 6 citation tones other than two \( Ru \) tones

According to the representations in (6), the two types of disyllabic sandhi are thus differentiated because one type disyllabic is initially \( Ru \)-toned underlyingly while the other type initially non \( Ru \)-toned underlyingly. With these words said, the representations of (6) are nevertheless still far from the reality of Wenzhou disyllabic sandhi. This is because Wenzhou disyllabic sandhi is actually a bi-directional process, meaning that both edges of disyllabics undergo sandhi. This is rather peculiar, as disyllabic sandhi in most of other Wu dialects is simply a one-edge phenomenon, as graphically depicted in Shanghai sandhi in (2).

However, the bi-directional sandhi in Wenzhou is by no means a balanced tone change on both edges. In particular, while the first syllable undergoes a drastic sandhi by pattern substitution, tonal change on the final syllable is fairly "mild", in the sense that the final sandhi is limited to register (i.e. Yin vs. Yan) neutralization with the basic melodic contour preserved. The situation is comparable to segmental merger of final

\( ^5 \) The exceptions to this tonal network are tones IIb and IIIb. Since both of them are of a Yan tone, they are supposed to share the same contour with their respective Yin counterparts IIA and IIIA. A tonal discrepancy like this has been observed across Chinese dialects, which can be accounted for from a diachronic point of view, but I will ignore the issue here because it would take me too far afield.
obstruent in German. Just like word-final obstruents in German lose voicing contrast, allowing the voiceless member only, disyllabic final tones in Wenzhou are suspended in Yin vs. Yan contrast, taking Yan tone only. Since register neutralization happens only on the right edge and, most importantly, independent of initial tonal change, I would like to term it right-edge neutralization for simplicity from now on, and accordingly modify the representations of (6) in (7).

(7) disyllabic tone sandhi in Wenzhou

<table>
<thead>
<tr>
<th>input</th>
<th>output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I: TR&lt;sub&gt;Ru&lt;/sub&gt; + T --- &gt; 1 + T --- &gt; [1 + TR]&lt;sup&gt;R&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Type II: T&lt;sub&gt;non-Ru&lt;/sub&gt; + T --- &gt; 42/53 + T --- &gt; [42/53 + TR]&lt;sup&gt;R&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

T = citation tone; TR<sub>Ru</sub> = 213 or 212
T<sub>non-Ru</sub> = the remaining 6 citation tones other than two Ru tones
TR<sub>R</sub> = right-edge neutralization (i.e., Yan tone).

To illustrate the foregoing expositions in some detail, I tabulate a partial list of Wenzhou disyllabic combinations in (8) below:

(8) Illustrations of Wenzhou sandhi

<table>
<thead>
<tr>
<th>1&lt;sup&gt;st&lt;/sup&gt; syllable</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt; syllable</th>
<th>II 'shang' Ascending</th>
<th>III 'qu' Departing</th>
</tr>
</thead>
<tbody>
<tr>
<td>I 'ping'</td>
<td>Ia 45</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ib 33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II 'shang'</td>
<td>IIa 45</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IIb 34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III 'qu'</td>
<td>IIIa 43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IIIb 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV 'ru'</td>
<td>IVa 313</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IVb 212</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In the table above the citation tones of the first and the second syllables are listed on the left second column and the top second row respectively. The sandhi tones of such combinations are given in the boxes where the two citation tones intersect. Thus, for example, a sequence of /34-45/ (i.e. tone IIb + tone IIa) surfaces up as [42-34], and /313-45/ (i.e. IVa + IIa) as [1-34]. Statistically, there are 16 possible disyllabic combinations framed within the column of Shang category, but they end up in two initial sandhi forms only. In particular, two distinct Ru tones (i.e. 313, and 212) surface as a low-level short tone 1. This is what I called Type I sandhi. The remaining 6 distinct initial tones are replaced by one high falling tone, namely 42. This is an instance of Type II sandhi. All the final tones, regardless of whatever initial sandhi is, take up the form of IIa tone (i.e. 34) with an understanding that IIa and IIb are merged because of the suspension of Yin vs. Yan on the right edge. As a further example, consider the right-most column of the table in (8), which frames all possible disyllabic combinations with their final syllable being underlyingly Qu-toned. Because of right-edge neutralization, the final Qu tones are merged as IIIb tone (i.e. 11), in a manner exactly like what has been seen with the final Shang tones. The initial Ru tones within this category surface up again as a low-level short tone, but the remaining 6 initial citation tones are replaced by another form of high falling tone, which is transcribed as 53. The difference between 42 and 53 is obviously due to the tonal difference in the final syllable.

To sum up, Wenzhou disyllabic sandhi has two patterns according to the initial sandhi forms as represented in (6), which is now reproduced here in (9a). The resultant two types of disyllabics carry a concomitantly different stress contour, as marked out in (9b). Specifically, Type I disyllabic demonstrates an iambic rhythm while type II a trochaic beat. The stress difference between type I and II can be accounted for an observation in Zheng-zhang (1964). That is, a phonetic emphasis on the initial syllable will be allowed with a Type II disyllabic, but not with a Type I syllable. In the latter case, an initial accentuation will disrupt the disyllabic combination, and destroy the sandhi effect with the initial syllable reverting to its original lexical tone.
Resolving Tension between Tone and Stress

(9) a. Wenzhou disyllabic sandhi patterns

Type I: $T^{Ru} + T \rightarrow 1 + T \rightarrow \{1 + T^R\}$

Type II: $T^{non-Ru} + T \rightarrow 42/53 + T \rightarrow \{42/53 + T^R\}$

$T$ = citation tone; $T^{Ru} = 213$ or $212$

$T^{non-Ru} = \text{the remaining 6 citation tones other than two } Ru \text{ tones}$

$T^R = \text{right-edge neutralization (i.e. Yan tone).}$

b. i). Type I sandhi and iambic contour

$\frac{T^R}{Ru} + \frac{T}{/} \rightarrow \{1 + T^R\}$

Ex. mu$^{212}$ + di$^{15}$ \rightarrow \{mu$^1$·di$^{34}$\} "sole"

ii). Type II sandhi and trochaic contour

$\frac{T^{non-Ru}}{Ru} + \frac{T}{/} \rightarrow \{42/53 + T^R\}$

Ex. vu$^{34}$ + sa$^{45}$ \rightarrow \{vu$^{42}$·sa$^{34}$\} "umbrella"

(* = stressed; . = stressless)

The schematic representations with illustrations in (9b) complete a survey of Wenzhou disyllabic sandhi with regard to surface word stress contour, but at the same time poses a clear-cut problem for metrical analysis: why is there a prosodic tension between sandhi and stress over Type II disyllabics, but not over Type I. In particular, over the initially Ru-toned combinations (i.e. (9bi)), which account for only a minor portion of Wenzhou disyllabics, tone sandhi and stress assignment work in a perfect tango. On the contrary, in the initially non Ru-toned combinations (i.e. (9bii)), which represent the majority of Wenzhou disyllabics, there is a clash between tone sandhi and metrification. In what follows I am proposing a resolution of the clash with the help of the current generative insights in stress phonology.

3. Deriving Disyllabic Metrical Structure

As a starting point, I assume that, following Halle & Vergnaud 1987 (H&V henceforth), the metrical grid is not directly built on syllables, but on a line of stress bearing units projected from the skeletal tier. This line of units, or base line, is termed line 0 within their framework. I will follow this
practice in this writing. Thus, while graphically sitting directly on the top of syllable symbol, line 0 asterisks in (10) are represented with the understanding that they are the projection of syllable nucleus.

\[
\begin{array}{c}
\sigma_1^T \sigma_2^T \\
T = \text{citation tone}
\end{array}
\]

In light of current moraic theory, what the representation of (10) says then is that the syllable should be underlyingly light in Wenzhou Chinese. A justification for this assumption is that, as noted earlier in the introduction, Wenzhou has neither diphthongs nor simple long vowels. In other words, the syllable nucleus in Wenzhou is underlyingly non-branching. Though empirical findings on this particular observation are not available in the literature at this point, this need not detain my argument here: the syllable rime in the language is inherently light. At least two lines of evidence can be furnished here. First, Wenzhou vowels lack a contrast in length. This can be seen by examining how the language behaves under the force of word minimality (McCarthy & Prince, 1995). For example, a CV monosyllable, when standing alone, is always pronounced with extended duration. This is obviously a strategy to meet word minimality moraically by lengthening the rime vowel to make it bimoraic (11a). Otherwise, the monosyllable tends to be headed by a semantically empty monosyllable (11bii), or to be reduplicated (11bii) to yield a disyllabic. Both (1bi) and (11bii) represent a strategy to meet word minimality at syllabic level.

\[
\begin{array}{c}
\text{(11) Word minimality effects in Wenzhou} \\
a) \quad \text{Word minimality at moraic level} \\
\text{pa --- > pa: "dad"} \\
\text{.μ --- > .μ: "uncle"}
\end{array}
\]

\[
\begin{array}{c}
b) \quad \text{Word minimality at syllabic level} \\
i) \quad \text{semantically empty lead} \\
\text{pa --- > a pa "dad"}
\end{array}
\]

There have been some findings on some better-known Wu dialects, e.g. Duanmu (1993) on Shanghai, and Yuan (1989) on Suzhou. These data indicate the rime in those dialects is underlyingly simple.
With the syllable nucleus said, what about the coda then? Like many other Chinese dialects, Wenzhou has only one type of complex rime: it is [VN]. Importantly, it is not entirely clear as to the phonological identity of the syllable-final nasal. For one thing, there is no minimal N-coda contrast in the language, which has been presumed to give rise to a heavy accent for Wenzhou natives to speak standard Mandarin, which has minimal N-coda contrast between [Vn] and [Vh]. On the other hand, a number of Wu dialects have been reported as having positional alternation between [VN] and nasalized [V] (e.g. Yuan, 1989). In particular, the [VN] is heard only when the word occurs alone, and [V] elsewhere. This can be explained away if we assume that the rime [VN] should be phonologically treated as [V], rather than a [V] plus a [N]. It is to meet word minimality requirement that the nasal surfaces as a coda in word final position. While further empirical work is needed to substantiate the argument above, we might alternatively as well assume that Weight-by-Position does not operate in Wenzhou. This assumption enables us to leave the syllable-final nasal outside of the rime. A comparable situation has been argued for word-final consonant in English. Similarly, [VC] counts as a complex rime and heavy in Latin, but as light in Lardil. All these observations point up one fact that moraic structure can vary from language to language (Hayes, 1989). Reasoning along this line, I assume that syllable-final nasal in Wenzhou is not a moraic segment.

In sum, syllable rime in Wenzhou is inherently monomoraic. This accounts for the base line representation of metrical projection in (10). It is fairly uncontroversial now among current metrical theories that heavy syllables bear inherent stress while light syllables get stress only by rule. Since all syllables in Wenzhou are light, then what is the rule for stress assignment? The rule I propose for computation of word stress contour in Wenzhou is of an iambic pattern. More specifically, the stress foot in Wenzhou is binary and right-headed, parsed from right to left.
According to Hayes's study (1995), a trochaic parse should be favored in a language like Wenzhou that lacks a quantitative opposition underlyingly. As indicated in (12), it seems not the case with Wenzhou, though. Within H&V framework, there are only two parameters relevant to foot parsing, that is, direction and headness. In other words, quantity does not have a role at least in initial foot parsing.\(^7\) Therefore, a choice between iambic and trochaic parses is mainly an empirical issue. One piece of empirical evidence for iambic parsing in Wenzhou can be seen from Type I disyllabics, over which, as illustrated in (9bi), tone sandhi synchronizes with metrification, resulting in "normal" word stress contour. Here are some more examples in (13).

\[\begin{align*}
\text{(12) Iambic prosody of Wenzhou} \\
* & \text{ line 1} \\
(* *) & \text{ line 0} \\
\sigma_1^L & \sigma_2^L
\end{align*}\]

The examples shows that a disyllabic starts with two full tones and then ends up in loss of the initial Ru tone to a low-level short tone. This tone sandhi pattern is coupled closely with iambic stress contour, which is graphically represented as a dot and a star respectively on the initial weak syllable and the final strong syllable. Translated into H&V's framework, the metrical derivation (13) can be spelled out in the following four steps:

\[\begin{align*}
\text{(13) Trochaic word contour of Type I disyllabics} \\
\text{pa}^{313} + \text{hu}^{33} & \rightarrow \text{pa}^1 \text{hu}^{33} \\
\text{"flowers"} \\
\text{ba}^{313} + \text{ho}^{33} & \rightarrow \text{ba}^1 \text{ho}^{33} \\
\text{"whitish shrimp"} \\
\text{ti}^{313} + \text{ly}^{11} & \rightarrow \text{ti}^1 \text{ly}^{11} \\
\text{"railway"} \\
\text{fi}^{212} + \text{de}^{11} & \rightarrow ^{*} \text{i}^{1} \text{de}^{11} \\
\text{"temperature"}
\end{align*}\]

(Data are from Zheng-zhang 1964b)

\(^7\) As pointed out by Hayes (1995: 76), in H & V's model "the distinction between quantity-sensitivity and quantity-insensitivity feet is not specified in the foot construction rules per se."
Resolving Tension between Tone and Stress

Metrical derivations of Type I disyllabic

\[
\begin{align*}
\text{line 2} & & * & * & * \\
\text{line 1} & & * & & (*)
\end{align*}
\]

\[
\begin{align*}
\text{line 0} & & & & (* & * & (* & *) & (* & *) \\
\sigma_1^R & \sigma_2^T & \text{--- > } & \sigma_1^R & \sigma_2^T & \text{--- > } & \sigma_1^S & \sigma_2^T & \text{--- > } & \sigma_1^S \\
\sigma_2^T & \text{--- > }
\end{align*}
\]

(input) Baseline projection Iambic parse from R to L End Stress (left)

Line conflation
1

3

4

(* * *)

\[
\begin{align*}
\sigma_1^S & \sigma_2^{TR} \\
\text{(output)}
\end{align*}
\]

\[
\begin{align*}
T & = \text{citation tone;} \\
TR & = \text{right-edge neutralization (i.e. suspension of Yin vs. Yan)} \\
R & = \text{Ru tone;} \\
S & = \text{sandhi tone (i.e. low-level short tone);} \\
R & = \text{Right} \\
L & = \text{Left}
\end{align*}
\]

The scenario in (14) is based on consideration of Type I disyllabics. However, I take one step further, and assume, on the grounds that Wenzhou disyllabics essentially follows a last-dominant paradigm of sandhi, that the iambic scenario in (14) is a generalized algorithm, whether they are of Type I or Type II.

The question then is: how come Type II disyllabics, as already showed in (9bii), will end up in a trochaic contour instead of the expected iambic rhythm? This is due to an interaction between tone and stress in the derivation of those disyllabic combinations, which is however not seen in the derivation of Type I disyllabics in (14). As noted earlier, the non Ru-toned initial syllable in Type II combinations surfaces up with a high falling tone (i.e. 42, or 53). The sandhi tone brings in its wake a lengthening effect on the rime vowel. There have been some works in the literature on the correlation between tone and vowel duration (e.g. Gandour, 1977; Maddieson, 1978). The basic observation is this: long vowels have developed under high falling tones, and/or high level tones,
regardless of vowel quality. This observation has been arrived at mainly from a diachronic point of view in Asian languages. The present data on Wenzhou disyllabic sandhi substantiates this insight from a synchronic point of view by appealing to a moraic analysis in (15).

(15) Vowel lengthening of high falling sandhi tone

\[
\begin{array}{c}
\text{-cons} & -\text{cons} & 42/53 & T \\
\mu & \mu & \mu & / \\
\end{array}
\]

\(42/53 = \text{high falling sandhi tone}\)

\(T = \text{citation tone}\)

\(\sigma_2 = \text{final syllable of disyllabics}\)

From the point of view of metrical grid theory as proposed by H & V, the additional \(\mu\) in (15) will necessarily project a new asterisk into the base line, which in effect deforms the metrical constituent structure established by original metrification.

(16) Deformation of metrical structure

\[
\begin{array}{c}
* \\
(\ast \ast) \\
\sigma_1^{\text{non-Ru}} \quad \sigma_2^T \\
\end{array}
\rightarrow
\begin{array}{c}
* \\
(\ast \ast) \\
\sigma_1^S \quad \sigma_2^T \\
\end{array}
\]

\(T = \text{citation tone}\)

\(\sigma_1^{\text{non-Ru}} = \text{non Ru toned syllable}\)

\(S = \text{high falling sandhi tone } 42 \text{ or } 53\)

Obviously, the resultant representation in (16) is illicit. For one thing, derived vowel lengthening destroys the originally parsed foot by adding a third mora to the binary structure. For another thing, iambic prosody does not permit the grouping of a heavy syllable (H) with a light syllable (L) as its dependent. That is, a (H 'L) grouping is prohibited for the reasons of quantity affinity inherent in iambic prosody (Hayes, 1995). To sum up, under the combined force of foot binarity and iambic quantity affinity, the deformed metrical constituent is then destroyed for
Resolving Tension between Tone and Stress

reconstruction. For technical convenience, I call this process iambic lengthening.

(17) Iambic lengthening

\[
\begin{array}{ccc}
\sigma_1^{\text{non-Ru}} & \sigma^T & \sigma^S \\
(\ast \ast) & \longrightarrow & (\ast \ast \ast) & \longrightarrow & \ast \ast \ast & \longrightarrow & \ldots \\
\end{array}
\]

The term iambic lengthening can be employed to cover a variety of situations, where iambic prosody exercises its organizing power in different stages of metrical derivations. Most appropriately, the term describes a situation in which a light syllable, while parsed as head of an iamb, becomes heavy through vocalic lengthening or consonantal germination. There is no lack of data in the literature that can be interpreted in this light. Take the Muskogean language Choctaw for example. The language is reported to have a vowel-lengthening rule, which targets every even-numbered short vowel except in the final position (Nicklas, 1975: 242-43).

(18) Iambic lengthening in Choctaw

\[(V^1 V^2) (V^3 V^4) \ldots \quad <V^n> \quad \longrightarrow \quad (V^1 V^2 : ) (V^3 V^4 : ) \ldots \quad <V^n>\]

As indicated in (18), the binary iambic parsing is enhanced by vocalic lengthening, which turns every instance of (L L) into (L H). From the metrical point of view, this segmental rule should really be termed iambic lengthening.

While iambic lengthening in Choctaw is induced by metrical parsing, just to reinforce its iambic prosody, the term might be also used to describe a mirror situation, wherein it is segmental lengthening that
Shane Xuexin Cao

derives iambic parsing. Take Yidin\textsuperscript{y} for example. One of the primary sources of length in the language is known as penultimate vocalic lengthening, which is applicable to words with odd number of syllables. It has been observed that vocalic lengthening has a dramatic effect on metrical parsing in the language. That is, when a long vowel falls in an even-numbered syllable, the word will have its even-numbered syllable stressed; otherwise the odd-numbered syllables are stressed (Dixon, 1977, cited in Hayes, 1982). Without going into details of his sophisticated reanalysis of this peculiar stress parsing dichotomy in Yidin\textsuperscript{y}, Hayes' (1982) insight into the phenomenon can be reinterpreted here as a different type of iambic lengthening, which derives the choice between iambic vs. trochaic parsing.

\begin{equation}
\text{guda:ga} \quad \rightarrow \quad (\text{guda:}) \, (ga)
\end{equation}

\begin{equation}
\text{wu\-naba:di\-n} \quad \rightarrow \quad (\text{wuna}) \, (ba:di\-n)
\end{equation}

(Examples are based on Hayes 1982: 105)

Returning to Wenzhou disyllabic sandhi in (15) through (17), I find what I have called iambic lengthening fits neither of the situations illustrated respectively by Choctaw and Yidin\textsuperscript{y}. On the one hand, the quantitative enhancement in Wenzhou happens not with the head of an iamb, like that in Choctaw. But rather it is with the non-head of the iamb. On the other hand, iambic lengthening in Wenzhou leads to the destruction of the original iambic metrical structure, unlike that in Yidin\textsuperscript{y}, where segmental quantitative enhancement gives rise to iambic parsing. As indicated in (15), tone sandhi upheaval occasions vocalic lengthening in Wenzhou disyllabics, which in turns deforms the underlying iambic prosody in (16), leading to the destruction of the original parsed constituent in (17) for metrical reconstruction. It is in this sense that iambic lengthening in Wenzhou serves as an interface for the interaction between tone sandhi and metrical structuring.
Resolving Tension between Tone and Stress

(20) Metrical reconstruction

\[ (* *) \rightarrow (* * *) \rightarrow * * * \rightarrow \ldots \rightarrow (* *) (*) \]

\( \sigma_{1}^{\text{non-Ru}} \quad \sigma^{T} \quad \sigma^{S} \quad \sigma^{T} \quad \sigma^{S} \quad \sigma^{T} \quad \sigma^{S} \quad \sigma^{T} \)

1). input 2). sandhi upheaval 3). metrical destruction refooting

\( \sigma_{1}^{\text{non-Ru}} = \text{non Ru toned syllable; } S = \text{high falling sandhi tone} \)

As implied by the derivational hiatus in (20), metrical reconstruction, however, does not proceed on a metrical clean slate. In particular, refooting does not begin right after metrical destruction. In between the two steps there is an accenting rule, which marks off the boundary of the heavy syllable by placing an opening bracket "(" in front of the appropriate asterisk on line 0. According to the metrical theory proposed by Halle (1990) and H & V (1987), an accenting rule can operate in both lexical and derived environments. In the Wenzhou case, the accenting rule comes to function upon derived iambic lengthening, hence the term iambic accenting in (21).

(21). Iambic accenting

\[ \ldots \rightarrow * * * \rightarrow \ldots \rightarrow (* *) (*) \]

\( \sigma^{S} \quad \sigma^{L} \quad \sigma^{S} \quad \sigma^{L} \quad \sigma^{S} \quad \sigma^{T} \)

This accenting bracket must be respected by the metrical parse to guarantee that the metrical parsing does not disrupt syllabic bracketing, meaning that a heavy syllable always constitutes a foot on its own. Otherwise, refooting which, exactly like the initial parse, follows a right-to-left binary right-headed paradigm,\(^9\) would split the heavy syllable by gathering two asterisks from the right edge into one foot.

Obviously, the immediate output of refooting, however, cannot be prosodically licensed for stress clash on line 1. The problem is then

\[^9\text{It can also be argued on the grounds of learnability that it is desirable for refooting to follow the same algorithm as does the initial parsing.}\]
resolved by a stress retraction rule, which pushes the left most asterisk to the next available position that is supported by a line 0 asterisk.

(22) Stress retraction

\[
\begin{align*}
 & \quad \ast \ast \quad \rightarrow \quad \ast \ast \\
\delta^S & \quad \delta^T \\
\end{align*}
\]

The rest of metrical derivations to follow are steps 3 and 4 as spelled out in (15), namely End stress and Line conflation:

(23) End stress (left) and Line Conflation

\[
\begin{align*}
 & \quad \ast \ast \quad \rightarrow \quad \ast \ast \ast \\
\sigma^S & \quad \sigma^T \\
\end{align*}
\]

In H & V’s work, Line conflation is motivated essentially for wiping out the secondary stress, which is case with Wenzhou disyllabic. On top of this, an additional function may be postulated here for Line conflation in a tonal language like Wenzhou: it wraps up disyllabic tone sandhi by trigging right-edge neutralization, namely, the suspension of tonal register contrast in Yin vs. Yan.

(24) right-edge neutralization

\[
\begin{align*}
 & \quad \ast \ast \quad \rightarrow \quad \ast \ast \ast \\
\sigma^S & \quad \sigma^T \\
\end{align*}
\]

To conclude the section, I summarize the metrical derivations of Type II disyllabics in (25a), tidying up all the details of how tone sandhi and metrification come to interact with each other in a two-way mode. The whole process is illustrated in (25b) in the form of νυυ24 sa35 (umbrella), which has been showed in (9bii).
Resolving Tension between Tone and Stress

(25)a.i. Mora is stress-bearing unit, which is projected as a line 0 asterisk.
ii. Line 0 parameter settings: binary, right-headed, right-to-left.
iii. Construct metrical constituent on line 0 by copying the head onto line 1.
iv. The initial metrically weak syllable loses its base tone to a high falling tone
   42 or 53, whichever is chosen by the base tone of the last metrically strong syllable.
v. Iambic lengthening by high falling sandhi tone deforms the existing metrical constituent, trigging a metrical reconstruction.
vi. Iambic accenting to bracket a heavy syllable as a foot of its own.
viii. Resolve stress clashing on line 1 by retracting the first asterisk to the next available left position supported by a line 0 asterisk.
ix. Line 1 parameter settings: unbounded, left-headed.
x. Construct constituent on line 1 and copy the head onto line 2, completing End stress.
xii. Conflate line 2 and line 1 to wipe out the last degenerate stress foot
xiii. Right-edge register neutralization on the last syllable tone.

--- >
baseline projection    \[**\]
baseline projection    \[**\]
iambic accented
Baseline projection    \[**\]
(iambic accented)
Baseline projection    \[**\]
(iambic accented)
Baseline projection    \[**\]
(iambic accented)
Baseline projection    \[**\]
(iambic accented)
Baseline projection    \[**\]
(iambic accented)
Baseline projection    \[**\]
(iambic accented)
Baseline projection    \[**\]
(iambic accented)
4. Conclusions

This paper is a case study of the metrical analysis of disyllabic sandhi in Wenzhou Chinese. Descriptively, there are two types of disyllabic sandhi according to the pattern substitution on the initial syllable. Type I refers to those disyllabics whose initial Ru tones are replaced by a low-level short sandhi tone. Type II is of those disyllabics whose initial non-Ru tones are replaced by a high falling tone 42 or 53, depending on the tonal category of the following syllable. In contrast with drastic tonal upheaval on the initial syllable, the final syllable of both types undergoes a "mild" sandhi, -mild because the sandhi is limited to register suspension with the melodic contour preserved. In this sense, Wenzhou disyllabic essentially follows a last-dominant paradigm. As attested in many Chinese dialects (Yip, 1980), the metrical reflex of last-dominant sandhi is an iambic parsing, which is substantiated in Wenzhou Chinese by its Type I disyllabics.

However, there is a prosodic mismatch between sandhi and stress over Type II disyllabics whose initial syllable is underlyingly non Ru-toned. In particular, while following last-dominant sandhi paradigm, this type of disyllabics, which represents the overwhelming majority of all disyllabic combinations in the language, surfaces with a trochaic, instead of iambic, contour. To resolve this prosodic tension, the paper proposes a two-way model of metrical analysis of tone sandhi, based on the framework developed mainly in H & V (1987). Specifically, just like Type I disyllabic, Type II disyllabics follow an iambic algorithm. However, during the metrical derivation, initial sandhi upheaval effected by a high falling tone substitution brings in its wake vocalic lengthening, a process I dub iambic lengthening for technical convenience. Under the combined force of foot binarity and quantity affinity inherent in iambic prosody, the iambic lengthening leads to refooting of the deformed metrical constituent. Remetrification then follows the exactly same iambic parameter settings as the original parse, but is critically preceded by an iambic accenting, which is to guarantee that the derived heavy syllable constitutes an independent foot on its own.
Resolving Tension between Tone and Stress

References

Shane Xuexin Cao
Macao Polytechnic Institute
e-mail: staneco@ipm.edu.mo