INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book. These are also available as one exposure on a standard 35mm slide or as a 17" x 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600
Tone and tone sandhi in Chinese

Zhang, Zheng-sheng, Ph.D.
The Ohio State University, 1988
TONE AND TONE SANDHI IN CHINESE

DISSERTATION

Presented in Partial Fulfillment of the Requirement for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By
Zheng-sheng Zhang, B.A., M.A.

The Ohio State University
1988

Dissertation Committee:
Mary E. Beckman
Marjorie K.M. Chan
Arnold M. Zwicky

Approved by
Adviser
Department of Linguistics
To My Family
ACKNOWLEDGMENTS

I am indebted to Arnold M. Zwicky for his guidance both in the dissertation project and in my graduate education. His dedication to linguistics and linguistic education has been inspirational. His continuous encouragement is instrumental in bringing the present work to completion. I am grateful to other members of my dissertation committee, who have given me many valuable suggestions concerning the style and content of the present work. Mary E. Beckman has provided me with keen insights and is always willing to help. Marjorie K. M. Chan has given me generous support, both in supplying me with materials on Chinese linguistics and in devoting time to discussions and comments.

To friends and family far and near, I am grateful for their moral support throughout the years of my graduate studies.
VITA

July 1, 1956 .................... Born - Shenyang, China

1982 ........................... B.A. Foreign Languages, Beijing Teachers' College, Beijing, China

1985 ........................... M.A. Linguistics Department Ohio State University, Columbus, Ohio

PUBLICATIONS


FIELDS OF STUDY

Tonal phonology, Morphology, Syntax, Pragmatics, Semantics
# TABLE OF CONTENTS

DEDICATION .......................................................... ii  
ACKNOWLEDGMENTS ..................................................... iii  
VITA ......................................................................... iv  

CHAPTER I. INTRODUCTION. ........................................... 1  

CHAPTER II. MANDARIN AS A SYLLABLE TONE LANGUAGE. ...... 7  

2.1. Mandarin tones ..................................................... 9  
2.2. Unit of representation ............................................ 11  
2.2.1. Separability of two types of arguments ................. 12  
2.2.2. Arguments for independent tiers ....................... 17  
2.2.3. Arguments for domain of tones ......................... 21  
2.3. One-to-one correspondence in Mandarin .................. 24  
2.3.1. Contour tones in polysyllabic morphemes ............ 26  
2.3.2. Against Exceptional Pre-linking ....................... 30  
2.3.3. The absence of positive arguments ................... 32  
2.4. Syllable as the domain of tones ........................... 34  
2.4.1. The proposal ............................................... 34  
2.4.2. Syllable tone and the OCP ............................... 37  
2.4.3. Syllable tone and the power of the theory .......... 39  
2.5. The issue of monosyllabicity ............................... 40  

Notes to chapter II. .................................................. 44  

CHAPTER III. TONE SANDHI IN MANDARIN ....................... 46  

3.1. Morphological tonal changes .................................. 49  
3.1.1. yi, qi, ba, bu ............................................ 49  
3.1.2. H tone in reduplication ................................ 52  
3.2. Purely metrically conditioned tonal rules ................ 55  
3.2.1. The Half Third Tone Rule ............................... 58  
3.2.2. The underlying form of the third tone ............... 60  
3.2.3. The Half Fourth Tone Rule ............................. 64  
3.3. The Second-to-First tone rule .............................. 66  
3.3.1. Problems with the statement of the rule ........... 68  
3.3.2. Stress patterns .......................................... 70  
3.3.3. Yip's statement of the rule ........................... 72  
3.3.4. Prosodic domains ........................................ 78  
3.3.4.1. Lenition and resyllabification .................... 78  
3.3.4.2. Domains for the Third Tone Rule ............... 81
3.3.5. Restatement ........................................... 83
3.4. The Third Tone Rule ..................................... 83
3.4.1. The problem ........................................... 83
3.4.2. The domain of Third Tone Rule ....................... 85
3.4.2.1. Linear versus hierarchical ......................... 85
3.4.2.1.1. Cheng (1973) .................................. 85
3.4.2.1.2. The necessity of cyclic application ............ 87
3.4.2.2. Syntactic versus prosodic .......................... 90
3.4.2.2.1. Kaisse (1985) .................................. 90
3.4.2.2.2. Liu (1980) ...................................... 91
3.4.2.2.3. Shih (1986) .................................... 91
3.4.2.2.4. The necessity of independent prosodic structures .......................... 94
3.4.3. The Present Analysis .................................. 95
3.4.3.1. The relevance of pragmatics and discourse contexts ........................................... 96
3.4.3.2. Lexical foot formation ............................. 106
3.4.3.3. Revised Foot Formation Rule ......................... 109
3.4.3.4. The relevance of the Branch Condition .......... 109
3.4.3.5. The necessity of reanalysis ......................... 115
3.4.3.5.1. Conflict between DM and IC ..................... 115
3.4.3.5.2. Analysis of superfoot .......................... 117
3.4.3.6. The role of prosody, morphosyntax and discourse ........................................... 119

Notes to chapter III ........................................... 121

CHAPTER IV. DOMAIN OF TONES IN DANYANG .................. 123

4.1. Melodies in Danyang ..................................... 124
4.1.1. Historical tonal categories ........................... 124
4.1.2. Sensitivity to historical categories ................ 126
4.1.3. Danyang tones ......................................... 127
4.1.3.1. Monosyllabic forms ................................ 127
4.1.3.2. Combination forms .................................. 129
4.1.4. Relationship between monosyllables and disyllables ........................................... 130
4.1.5. Insufficiency of historical explanation ............ 131
4.1.6. Restriction of melodies ................................ 135
4.1.7. Present analysis ....................................... 136
4.1.7.1. Inventory of melodies ............................... 137
4.1.7.2. Monosyllabic forms .................................. 137
4.2. Domain of tones in Danyang ............................. 138
4.2.1. Phonological units as domain ....................... 139
4.2.2. Prosodic structure in Danyang ...................... 141
4.2.2.1. Trisyllabic forms .................................. 142
4.2.2.2. Prosodic structures in names and transliterations ........................................... 148
4.2.2.3. Quadrisyllabic forms ............................... 151
4.2.2.4. Stress patterns .................................... 157
4.2.3. Trisyllabic and quadrisyllabic melodies ............. 158

vi
CHAPTER I

INTRODUCTION

In this dissertation, I examine three Chinese languages, Mandarin, Danyang and Shanghai, with particular attention to the following two issues: first, the units of tonal representation appropriate for these languages; and second, the domains in which tones interact, i.e., the domain of tone sandhi. The first issue is a long-standing one and discussions of it are found in the works of Pike (1948), Welmers (1962), McCawley (1968) and Leben (1973). The second concern of the present thesis arises out of the recent interest in prosodic phonology (Hayes, 1984; Nespor & Vogel 1986) and the interface between phonology, morphology and syntax, represented by works of Chen (1985, 1988), Kaisse (1985), Pullum & Zwicky (forthcoming), Selkirk (1984, 1986) and Zwicky and Pullum (1983) and Zwicky (1985).

The present work defends the position that the unit of tonal representation differs from language to language. It is argued that it is the syllable or the
segment that is the unit of tonal representation in Mandarin (chapter 2); for Danyang it is the foot (chapter 4), and finally, the unit of tonal representation is the phonological word in Shanghai (chapter 5).

The implication of the conclusion above is that melodies on polysyllabic forms cannot always be a function of the component syllable tones but must be lexically associated with their melody-bearing units. This is amply demonstrated in chapter 4 and chapter 5, where Danyang and Shanghai melodies are discussed.

The present work demonstrates that hierarchical phonological structure is relevant to tonal phenomena in various ways. The hierarchical phonological structure here includes strength relationship in the sense of metrical phonology (Liberman and Prince 1977; Halle and Vergnaud 1978; Hayes 1981; Selkirk 1980, 1984) and prosodic domains in the sense of prosodic phonology (Nespor and Vogel 1986). It will be shown that tonal rules previously formulated quite unrevealingly in purely tonal terms turn out to be metrically conditioned. The various half tone rules in Mandarin belong to this type, as I will show in section 3.2 Reformulation making use of prosodic domain is also given, in section 3.3 to the Second-to-First tone rule, previously formulated in strictly tonal terms.
Prosodic domains non-isomorphic with morphosyntactic organization are found to serve as the domain of tone sandhi and other phonological phenomena. Specifically, the foot, which is argued to be the basic domain of the Third Tone Rule, is used in the fast speech Second-to-First Tone Rule and segmental lenition as well in Mandarin. The foot also serves as the domain of tonal melodies in Danyang.

In addition to the foot, larger prosodic units are also motivated by language-particular facts. In Mandarin, it is argued that Yip's metrical system for Mandarin, which posits only two levels, the foot and the phrase, is too restricted. In order to get the correct stress value assignment, at least an intermediate level between the foot and the phrase is needed. In the discussion of the Mandarin Third Tone Rule, it is found that Shih's (1986) Foot Formation Rule, which constructs the tone sandhi domain on the basis of only immediate constituency cuts and the rhythmic patterns of the language, is empirically inadequate. Higher prosodic constituents, delimited according to the discourse structure, must play a role in the account of Mandarin Third Tone Rule as well.

In both Danyang and Mandarin, a preference for left-branching prosodic structure is observed. The preference for left-branching structures is attributed
to the tendency toward left-dominance. If true, this will give a unified explanation for not only the prosodic structures observed and but also the transition from syllable tone to word tone. It is hypothesized that left-dominance first gives rise to left-branching prosodic sandhi domains, which then become melody-bearing units.

The variety of Mandarin discussed here is that spoken in the city of Beijing, on which the standard language Putuonghua is based. This fact is important because the tonal rules discussed here, for example, the H tone in reduplicated forms, and especially the fast speech Second-to-First tone rule, are special to Beijing Mandarin. The Mandarin-speaking area is very large, extending all the way to the northeast, northwest and southwestern parts of the country. Geographically, Beijing is located in the Eastern province of Hebei in the northern part of China.

Shanghai is language spoken in the metropolitan city of Shanghai, south of the Yangtse River and right on the Pacific coast. It belongs to the Wu language family spoken in the coastal provinces of Jiangsu and Zhejiang.

Danyang is spoken in the city of Danyang in Jiangsu province north of the Yangtse River. Geographically then, Danyang is located between Beijing and Shanghai,
being closer to the latter than the former. Linguistically, Danyang has been classified as a Wu language, although features of Mandarin are also evident in the language.

The organization of rest of the work is as follows.

In chapter II, I argue that tones in Mandarin are lexically associated with their pitch-bearing unit, the syllable. I propose to eliminate association conventions, without giving up the independence between the tonal and the segmental tiers.

In chapter III, three types of tonal rules in Mandarin are discussed. The first kind applies to specific morphemes and constructions. The second kind is metrically conditioned. The third kind, including the Second-to-First Tone Rule and the Third Tone Rule, apply in prosodic domains.

In chapter IV, Danyang is examined. It will be shown that the melody-bearing unit in Danyang is the foot; and melodies on feet cannot be derived from citation tones. The similarity between Mandarin and Danyang prosodic structures is also discussed.

In chapter V, the notion of word tone and criteria for identifying the word tone in the literature are examined. Then a word tone analysis of Shanghai is proposed. Arguments are given against deriving word
melodies from citation tones. Finally, it will be shown that the domain of melodies in Shanghai is the phonological word rather than the grammatical word.

Chapter VI puts forward the hypothesis that the tendency toward left-dominance is the motivation for both the change from right-dominance to left-dominance and also the transition from syllable tone to word tone languages. Left-dominance also gives a unified explanation to various facts in Mandarin and Danyang, including the way prosodic domains are formed.
CHAPTER II

MANDARIN AS A SYLLABLE-TONE LANGUAGE

The observation that Mandarin is a syllable-tone language is by no means new. It is customary to hear that every stressed syllable in Mandarin has one of the four lexical tones. More formally, Mandarin serves as the paradigm lexical tone language in Pike's definition of a tone language, 'a language having lexically significant, contrastive, but relative pitch on each syllable' (Pike, 1948: 3).

Why then do we have to say more on the matter? The reason for reopening the question is that in most versions of autosegmental theory, tones are represented on the morpheme. This usual representation is due, in large part, to Leben's (1973) argumentation, predating autosegmental phonology, for the suprasegmental representation of tones. In the context of Chinese, the morphemic representation of tones was adopted in Yip's (1980) dissertation, the first to study the tones of Chinese in the autosegmental framework.

It happens to be true to a large extent that the syllable and the morpheme are coterminous in Mandarin, so
that to take either as the unit of tonal representation gives the same result in most cases. But the question still remains: is it the syllable or the morpheme that is really the unit of tonal representation in Mandarin?

Another reason for asking the question is typological. In recent tonal literature on Chinese, a distinction between syllable tone and word tone has been made (Shih 1986). Mandarin is said to be a syllable tone language while Shanghai and some other Wu languages are said to be word tone languages. In having syllable tones, Mandarin will also differ from other word tone languages cross-linguistically. Mandarin will differ from Mende, the African word tone language studied by Leben (1973). To the extent that such typological distinctions can be made, the syllable tone status of Mandarin needs to be established explicitly.

In this section, I examine more closely the claim that Mandarin is a syllable tone language. I provide evidence for the claim and will also discuss the theoretical implications of syllable tone in the autosegmental theory of tone. I will show that the autosegmental theory of tone does not exclude the possibility of representing tones in the domain of the syllable or even the segment.
2.1. Mandarin tones.

Before discussion of the tones in Mandarin, a few words about their phonetic properties and their notation are in order.

Mandarin Chinese has four distinctive tones. Each of the stressed syllables in Mandarin has one of the four tones. The first tone is high level; the second tone is high-rising; the third tone is falling-rising; the fourth tone is high-falling. The following quadruplet has been used to illustrate the four tones of Mandarin:

<table>
<thead>
<tr>
<th></th>
<th>'mother'</th>
<th>'hemp'</th>
<th>'horse'</th>
<th>'curse'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>55</td>
<td>35</td>
<td>214</td>
<td>51</td>
</tr>
<tr>
<td>2</td>
<td>HH</td>
<td>LH</td>
<td>LLH</td>
<td>HL</td>
</tr>
</tbody>
</table>

Unstressed syllables have what is called the 'neutral tone', whose pitch height is not specified but depends on that of the preceding syllable. The pitch levels of the neutral tone after stressed syllables are as follows (Chao, 1968):

- half-low after 1st tone. Ex. 1 N ta de 'his'
- middle after 2nd tone. Ex. 2 N hwang de 'yellow'
- half-high after 3rd tone. Ex. 3 N ni de 'yours'
low after 4th tone. Ex. da de 'big'

The neutral tone is distinctive, because we can find minimal pairs contrasting it with the other tones.

Various notational devices are used to mark Chinese tones. Four most commonly used ones are given here. The single digits 1, 2, 3, 4 are used to represent the four tonal categories in Mandarin. Next, we find two ways to represent tone values in the Chinese tonal literature, both devised by Chao (1934). The pitch contours of tones are represented by digits, i.e., by specifying the pitch height of the beginning and the end points of the contours. The tonal range is divided into five levels, '5' being the highest and '1' the lowest. Corresponding to these numerical descriptions and below them are tone letters, which are stylized representations of contour types, with a vertical line representing the tone space and the stroke to the left of the line representing the contour. The last one, using tonal features of H and L, is used in the generative studies of Chinese tones. The particular representation given here is that of Yip (1980).

In the dissertation, unless specified otherwise, the following practice with regard to notation will be adopted. When tonal categories are referred to, single digit, Arabic numerals will be used. When tonal values
or melodies for these tonal categories are referred to, either Chao's system of tone numbers or H, L and M will be used. As the question of how to represent tones with a feature system is not central to this dissertation, the H, L and M should not be taken strictly as features.

The duration of the syllable in Mandarin varies depending largely on the tone it carries. Of the five tones, the neutral tone is the shortest. Its length was measured and shown to be a little less than half of that of syllables with other tones. (Zadoenko, cited in Lin 1963) The longest tone is the third tone in its citation form, the falling-rising tone. Of the rest of the three tones, there is disagreement as to their relative lengths. But for our purposes here, the relative length of the first, second and fourth tones is not important.

2.2. Unit of representation.

The question of where tonal features are represented has been answered in many different ways in the literature. We have, for example, the segmental position represented by Woo (1969), Maddieson (1971), Fromkin (1968); we also have Wang's proposal (1967) that the syllable is the domain of tonal features; more recently we find the suprasegmental representation of tone argued for by Leben and assumed now by almost all
autosegmental phonologists. In Leben's proposal, tone features are represented on the morpheme, separately from the segments, in the lexical representations. These tonal features are later mapped onto the segments.

Autosegmentalists, starting from Goldsmith (1976), continue representing tones on the morpheme and associate them to segments by association rules and conventions.

We thus observe an association between the autosegmental position and the morphemic representation of tones. However, as I will argue below, such an association is not a necessary one and that it is possible to have an autosegmental theory that represents tonal features on phonological units such as syllables or even segments, however contradictory in terms it may seem.

2.2.1. Separability of two types of arguments

Let me first make precise what the terms 'autosegmental', 'segmental' and 'suprasegmental' mean.

The term 'autosegmental' means that features can be represented on separate tiers. It characterizes the 'independence' between features of different kinds.

The term 'suprasegmental', in the context of discussions about the units of tonal representation, means larger than the segment, i.e. syllabic, morphemic etc. In Leben's dissertation (1973), entitled
Suprasegmental Phonology, the term is exclusively used to refer to the morphemic representation of tonal features.

The term 'segmental' has been taken to contrast with 'autosegmental' in the sense that in 'segmental' theories of Woo, Fromkin and Maddieson no independent tiers are posited. But 'segmental' has been used to contrast with 'suprasegmental' as well, for example in Leben (1973), in the sense that the segment is the domain of tonal features. Most often, when the term 'segmental' is used, the two senses are grouped together.

To avoid the ambiguity of the terms 'segmental' and 'suprasegmental', the present thesis chooses to use two different terms instead. The relevant distinctions are the independence between tiers on the one hand and the domain of tonal features on the other hand. In the first case, the term 'autosegmental' seems acceptable; but since I will be comparing theories from different time periods, a more neutral pair of terms 'multi-tier' versus 'single tier' will be used. In the second case, since I am not going to address the question of whether tonal features are represented on the syllable as opposed to the segment (for discussions from both sides, see Wang (1967) and Leben (1973)), I will group the two together as phonological domains in contrast to the morpheme, which is a morphological domain.

Another complication arises when we consider the
independence parameter in relation to the organization of the grammar. Tiers in some theories are independent throughout the derivation; in some other theories, they are never independent; in yet some other theories, they are independent phonetically but not underlyingly; the opposite holds in still another set of theories.

The following chart gives the eight logical possibilities resulting from the contrasts in domain and independence of features, with reference to some of the actual proposals in the literature.

<table>
<thead>
<tr>
<th>DOMAIN</th>
<th>PHONOLOGICAL</th>
<th>MORPHOLOGICAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>multi-tier</td>
<td>Goldsmith (1976)</td>
</tr>
<tr>
<td>N</td>
<td>throughout</td>
<td>Leben (1978)</td>
</tr>
<tr>
<td>D</td>
<td>multi-tier</td>
<td>Yip (1980)</td>
</tr>
<tr>
<td>E</td>
<td>(underlying)</td>
<td>Leben (1973)</td>
</tr>
<tr>
<td>P</td>
<td>multi-tier</td>
<td>Goldsmith (1976)</td>
</tr>
<tr>
<td>E</td>
<td>(phonetic)</td>
<td>(last chapter)</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>single-tier</td>
<td>Woo (1969)</td>
</tr>
<tr>
<td>E</td>
<td>throughout</td>
<td>Fromkin (1968)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maddieson (1971)</td>
</tr>
</tbody>
</table>

To represent the positions I am arguing for and against, I will for the sake of convenience use the labels 'syllabic representation of tone' and the 'morphemic representation of tones' respectively, without differentiating the syllable from the segment. By the
same token, although I will use the term 'syllable tone', whatever I say about it will also cover segment tones.

Most of the theories proposed fall into the top-right and bottom-left cells; that is to say, the favorite combinations seem to be single-tier theories with phonological units as the domain of tone features and thoroughly multi-tier theories with morphological units as the domain. The bottom-right cell is impossible since if the morpheme is the domain of tones a single-tier theory will not be able to accommodate the independent tonal features. Between the single-tier theories and the most full-fledged multi-tier theory of Goldsmith, Leben's theory (1973) deserves mention here. In his theory, the tonal features are represented separately in lexical entries from segmental features. It is therefore underlyingly multi-tiered. This much Leben's theory shares with Goldsmith's theory. The difference between Leben's theory and a full-fledged autosegmental theory like that of Goldsmith is that in the former tonal features are mapped onto segments eventually while in the latter the independence of tiers is maintained throughout the derivation.

What about the top-left box then, assuming that tiers stay independent throughout the derivation? Although few existing theories - Camhi (1985) is one - claim allegiance to this position, I claim that a fully
multi-tiered theory with the syllable or the segment as
the domain of tones is logically possible.

The position, in a nutshell, is illustrated below:

\[ T \quad T \quad T = \text{tone} \]
\[ \text{[ | | | ]}_m = \text{morpheme} \]
\[ s \quad s \quad s = \text{segment/syllable} \]

This picture looks neither outlandish nor surprising. After all, this is what is found in original autosegmental theory after tones are associated with segments. Independence between the tonal tier and the segmental tier is still maintained. The outstanding characteristic of the theory is that tones and tone-bearing units are linked already in the lexical representation, unlike most autosegmental theories. This means that although tones and tone-bearing units are on different tiers underlyingly, they are not separate from each other in lexical entries, as in the case of autosegmental theory of Goldsmith and the suprasegmental theory of Leben.

Given the logical possibility of such a position, I will now review some of the arguments for the separation of tiers and some of the arguments for representing tonal features on the morpheme and show that the two kinds of arguments are independent of each other. Therefore, the restriction of theories to the most favored combinations shown in the top-right and bottom-left cells in the chart
above have no particular empirical justifications.

2.2.2. Arguments for independent tiers. There are a number of familiar arguments first used by Goldsmith (1976) to motivate independent tiers: Stability, Contour Tone, Floating Tone and Tone Spreading. Yip (1980) used similar arguments to motivate the autosegmental analysis of Chinese tones. We will present Yip's arguments in turn.

The Stability argument utilizes the fact that when tones are deleted their segments can remain and vice versa. In Mandarin, for example, we have the following words:

<table>
<thead>
<tr>
<th>lao hu</th>
<th>xiao jie</th>
</tr>
</thead>
<tbody>
<tr>
<td>old tiger</td>
<td>small sister</td>
</tr>
<tr>
<td>'tiger'</td>
<td>'miss'</td>
</tr>
<tr>
<td>3 3</td>
<td>3 3</td>
</tr>
<tr>
<td>2 3</td>
<td>2 3</td>
</tr>
<tr>
<td>2 N</td>
<td>2 N</td>
</tr>
</tbody>
</table>

In these compounds, both morphemes have the 3rd tone underlyingly. The first changes to a 2nd tone due to the Third Tone Rule, which changes a 3rd tone to a 2nd tone when another 3rd tone follows. A tone deletion rule then applies to the second syllable, yielding a neutral tone on that syllable. The loss of tone in the second syllable leaves the segmental material intact. In Cantonese, the opposite happens:
When the segments of the second syllable yat\(^5\) are deleted, the high tone, designated as '5' here, remains and becomes part of the tonal contour of the first syllable.

The Contour Tone argument is based on the formal impossibility of having sequentially ordered feature specifications in the type of feature matrixes found in Sound Pattern of English (SPE). The following feature matrix is deemed impossible under the standard theory:

\[
\begin{array}{c|c}
\text{[high]} & \text{[-high]} \\
\text{ + back } & \\
\text{ +continuant } & \\
\end{array}
\]

The feature [high] is a tonal feature. The introduction of the independent tonal tier provides us with an alternative means of representation:

\[
\text{[+high] [-high]}
\]

\[
\begin{array}{c}
\text{[+back]} \\
\text{[+cont]}
\end{array}
\]

The Floating Tone argument can be demonstrated with an example from Cantonese. The diminutive of nouns can be formed by attaching a H tone. For example, the diminutive of the surname, Chen (in broad phonetic transcription):
The second syllable in this example changes from 21 to 35 in tonal value due to the attachment of the high tone. The diminutive morpheme then is the segmentless high tone. The representation of such morphemes will be problematic in theories without an independent tonal tier.

Finally, the Tone Spreading argument is demonstrated with examples from Shanghai by Yip. One such example is:

```
thy 'heaven, nature'  thy sang 'natural'
\/
53  | |  5 3
```

In Yip's analysis, the tone on the second part of the compound gets deleted in the process of compounding, and the second part of the tonal contour from the first syllable spreads rightward to cover the second syllable as well. The spreading analysis seems more insightful than a non-autosegmental analysis.

The above arguments can be called 'independence arguments'. As they are convincing arguments, I take the independence of tones from other features to be well-established.

But the establishment of the independence of tones still leaves open quite a few different options as to
what the theory will look like. To wit:

(a) An autosegmental theory like that of Goldsmith (1976), in which tones stay independent throughout the derivation. Tones are independent underlyingly, because they are represented separately from the segments in lexical representations; they are independent after being associated with tone-bearing units because rules can still refer to them alone; they are still independent at the phonetic level, as it is the intent of Goldsmith's (1976) theory to capture the lack of synchrony and the 'orchestra effect' in speech production. This kind of theory is located at the top-right cell of the chart in the last section.

(b) A suprasegmental theory like that of Leben (1973), in which tones start out independently---being represented separately in the lexical entry---and end up mapped onto segments. The theory is at least partially similar to an autosegmental theory mentioned above. It can be found in the top middle part of the chart.

(c) A theory entertained in the last chapter of Goldsmith's dissertation, 'The Autosegmental Index'. This theory seems to be a mirror image of Leben (1973) because it takes the autosegmentalization at the phonetic level as given and hypothesizes 'deautosegmentalization', collapsing autosegmental tiers into one tier, at the more abstract, lexical level. In our chart, it is located
below Leben (1973).

It is entirely within the realm of possibility then that tones are linked to tone-bearing units at the lexical level, unlike in most current versions of autosegmental theories. Tones may link to their tone-bearing units underlyingly and still exhibit autosegmental behavior in derived environments.

2.2.3. Arguments for domain of tones. The arguments used by Welmers (1962) and Leben (1973) for representing tonal features on the morpheme, rather than the segment/syllable, are of two types. The first type is based on the restriction of tonal melodies, and the second type has to do with the constant number of tonal melodies in combinations of different lengths.

The first argument goes like this: in languages like Mende, there is a constraint against combinations of HLH, which holds for words with an arbitrary number of syllables. The constraint can be stated only if tones are represented on domains larger than the syllable.

The second argument can be illustrated too with Mende data. In Mende, there are altogether five melodies: H, L, HL, LH, LHL. The melodies are mapped onto strings of different lengths. Take the last melody LHL for example. It can be mapped onto the following three words:
In the first case, we have all three tones associated with the same vowel; in the second case, we have the first L associated with the first vowel and the rest of the tones with the second vowel; in the last word, we have one-to-one association between tones and vowels. In the segmental theory, there would have to be four different tonemes: [LHL], [L], [H], and [HL], where the square brackets indicate syllable boundaries. Apart from the unrelatedness of the three sequences, we would also have to posit complex tonemes like LHL and HL for the language.

The two arguments are actually related and in a sense can be regarded as one. The second argument by itself does not have any force. Although Contour Tones is one argument for the autosegmental representation, it is not an argument for the morphemic representation of tones. If we allow many-to-one association between tones and segments, contour tones can be represented on the segment/syllable, albeit on a different tier. The relatedness of the melodies is also rather doubtful. They are related only in terms of the non-occurrence of the other melodies and not in terms of meaning. The three words in the Mende examples are not at all related.
semantically. And since they are simple words, the tonal melodies cannot be morphemes but must be tonemes capable of distinguishing the meaning of the morphemes.

So why choose the morpheme as the domain of such tonemes? Relatedness is apparent only when the number of melodies is small. Suppose that there are 14 melodies instead of 5 in combinations of up to three syllables - in other words, that the full range of possibilities is exemplified \((2^3 + 2^2 + 2^1 = 14,\) calculated with the formula of \(S = T^n,\) where \(S\) is the sum of melodies, \(T\) the number of tones and \(n\) the number of syllables):

\[
\begin{array}{ccc}
H & HH & HHH \\
L & HL & HHL \\
& LH & HLH \\
& LL & HLL \\
& & LHH \\
& & LHL \\
& & LLH \\
& & LLL \\
\end{array}
\]

We still find the five melodies, underlined above, but few people will now regard them as being related.

Both arguments used for the morphemic representation of tones derive their force from the fact that not all contrasts in syllables are realized so that the relevant unit to observe contrasts is not the syllable but the morpheme. We may then say that these are 'contrastive unit' arguments.

The 'independence' arguments and the 'contrastive
unit' arguments are independent of each other, because the former says nothing about the domain of tonal features while the latter makes no direct claims about the independence between tiers.

The separability of the tonal tier from the segmental tier by no means necessitates representing tones on the morpheme. Therefore, even with the autosegmental nature of tonal features firmly established, we do not have to accept without question the morphemic representation of tones.

2.3. One-to-one correspondence in Mandarin.

While accepting the autosegmental nature of the tonal tier, the present thesis challenges the position which represents tones as features of the morpheme. The first argument against representing tones on the morpheme in Mandarin is based on the observation that the correspondence between tones and tone-bearing units in Mandarin is one-to-one, which is what makes Mandarin a lexical tone language in the strict sense of Pike (1948) in the first place. This can be seen in the following examples:

1 4 2 4 3  3
sheng diao xue xiao ling dao
'tones' 'school' 'leader'

If tonal features were to be represented on the
morpheme, then it would really be a matter of accident that such one-to-one correspondence happens to hold. In any case, the morphemic representation theory predicts the lack of one-to-one correspondence as much as one-to-one correspondence.

It may be argued that since the syllable and the morpheme are to a large extent coterminous in Mandarin, the one-to-one correspondence observed is actually that between morphemes and tones and not between syllables and tones. Such correspondence can be accounted for by a morpheme structure condition to the effect that each morpheme has only one tone. Just looking at compounds and phrases, which constitute most of the polysyllabic expressions in Mandarin, we might agree that this is true. But the existence of polysyllabic morphemes, most of which are transliterations of foreign names, shows that this is not the case. For example:

1  4  3  
fei li pu  
'Philip'  

In examples like this, the one-to-one correspondence is positively between tones and syllables.

The observant reader will have noticed that I have used the unitary symbols 1, 4, 3 to represent the actual tonal contours H, HL and L respectively. The argument above then assumes the unitary analysis of contour tones.
Our general argument against representing tones on morphemes in Mandarin, however, does not depend on the particular theoretical stance taken on the question of contour tones. The argument can still be constructed assuming the sequential analysis of contour tones. The argument does not concern the missing of generalizations in one-to-one correspondence, but rather concerns the difficulty of aligning contour tones with syllables in polysyllabic morphemes.

2.3.1. Contour tones in polysyllabic morphemes. In the sequential analysis of contour tones, Mandarin is said to have the tonal melodies HH, LH, LL and HL underlyingly (Yip, 1980), corresponding to the 1, 2, 3, 4 in the unitary notation respectively.

For the sake of clarity, let me first present the unproblematic cases, which are tones on monosyllabic morphemes. If we take the morpheme as the domain of tonal assignment and furthermore assume cyclic tonal assignment, contour tones can be aligned to morphemes by a simple left-to-right, one-to-one mapping plus the operation of the Well-formedness Condition (WFC) postulated by Goldsmith (1976). The WFC, as stated by Goldsmith, is a condition on the permissible association between tones and tone-bearing units. It is quoted below (p.27):
Well-Formedness Condition (WFC):
(1) All vowels are associated with at least one tone.
All tones are associated with at least one vowel.
(2) Association lines do not cross.

An example of the unproblematic cases is the compound xue xi 'study':

\[
\begin{array}{cccccccc}
\text{LH} & \text{LH} & \rightarrow & \text{LH} & \text{LH} & \rightarrow & \text{LH} & \text{LH} \\
xue & xi & \rightarrow & xue & xi & \rightarrow & xue & xi \\
\end{array}
\]

Since the morpheme is to a large extent coterminal with the syllable in Mandarin, unproblematic cases like the above constitute the majority. But polysyllabic morphemes do exist; and they pose serious problems. To use the example fei li pu 'Philip' from the last section again, we see that the result in this case is even worse than the one-to-one correspondence observed before, as the tones cannot be mapped correctly onto the syllables, no matter which instantiation of the association principles is used and no matter whether Obligatory Contour Principle (OCP) is observed or not. The OCP, originally proposed by Leben (1973) and given the present name by Goldsmith (1976), is a constraint on the possible tonal melodies by which adjacent identical tonal segments are collapsed into one. For example, we can use the left-to-right association without obeying the OCP:
Observing the same association principle as well as the OCP, we get:

*[[H] [H L] [L] [H] [L] by OCP
  \____________/  \________ |
  fei li pu

Or, we can use right-to-left association, with or without the OCP:

*[[H] [H L] [L] [H] [L] by OCP
  / \____________/  \________ |
  fei li pu

*[[H] [H L] [L] [H] [L] by OCP
  \________ \____________/  \________ |
  fei li pu

We can even use edge-in association, as suggested by Yip (1988), which associates the left and right edges first and then fills in the middle by phonetic extrapolation or by default:

*[[H] [H L] [L] [H] [L] by OCP
  \________ \________ \____________/  \________ |
  fei li pu

Only when the contour tone is in the middle portion and when we do not assume the OCP, can we manage to get the
right association:

\[
\begin{array}{ccc}
\mid & \vee & \mid \\
fei & li & pu
\end{array}
\]

Similarly, such accidental success will be achieved in a number of combinations, all without assuming the OCP: (1) contour tone on the last syllable and left-to-right association; (2) contour tone on the first syllable and right-to-left association. But if there is more than one contour tones, none of the algorithms will work. As the limiting case, consider the following morpheme with three contour tones:

\[
\begin{array}{cccc}
L & H & L & H \\
\vee & \vee & \vee \\
pu & tao & ya
\end{array}
\]

'Portugal'

The problem, as we can see, lies in the inability to reconcile the distribution of contour tones in Mandarin with any sort of regular tonal assignment algorithm. There are potentially two ways to approach the problem, under a sequential interpretation of contour tones: (1) modifying existing tonal assignment algorithms; (2) treating the problematic cases as exceptional. Let us examine the two approaches in turn.

Modifying existing tonal assignment algorithms is unrealistic, because what the Mandarin facts call for is an extremely powerful algorithm. We have to be able to
get contour tones at any position and simple tones at any position as well. Such an algorithm will predict non-existing tonal configurations, for instance mapping all tones onto a particular syllable, leaving some syllables toneless.

2.3.2. Against Exceptional Pre-linking. Treating the problematic cases as exceptional is equally undesirable. But first let us see how exceptional cases are handled in a morphemic representation theory. A particular proposal is given in Leben (1978). Leben noticed that there are some contrasts in Mende that are not expressible with regular tone mapping conventions. One such contrast involves two words that are given below:

1) H L 2) *H L H L
   |    |    |
ngila ngongo ngongo

The first word can be derived straightforwardly with the regular left-to-right, one-to-one association. The second one, however, would be ill-formed if derived the same way. To ensure the correct form, the high tone is pre-linked to the second vowel, and the WFC operates to give the correct output. Notice in this particular case, positing a HHL melody and violating the OCP at the same time gives the same result, as pointed out by Odden (1986). Our present concern, however, is not with the
correct analysis of the word in question but rather the machinery of pre-linking.

The limiting case we had above would be derived as follows, assuming pre-linking:

\[
\begin{array}{ccc}
L & H & L \\
\text{pu} & \text{tao} & \text{ya} \\
\end{array}
\]

\[
\frac{\text{L H L H L H}}{[\,\checkmark \checkmark \checkmark \,]_m} \text{ lexical representation with pre-linking}
\]

\[
\begin{array}{ccc}
L & H & L \\
\text{pu} & \text{tao} & \text{ya} \\
\end{array}
\]

\[
\frac{\text{L H L H L H}}{[\,\checkmark \checkmark \checkmark \,]_m} \text{ mapping applies vacuously}
\]

But this is certainly not what we should be doing. Pre-linking is undesirable in the case of Mandarin, because the problematic cases are not exceptional at all. First of all, of the four melodies in Mandarin, H, LH, L, and HL, at least two are contour tones (assuming the third tone to be L rather than the more surfacy LLH). Second, contour tones in 'unexpected' positions, i.e., non-finally, are as common as those in expected positions, say at the end of the morpheme. Third, in the limiting case given above, where all the syllables are linked to contour tones, the regular association conventions become vacuous.

Treating a large part of contour tones as exceptional not only misrepresents the language but also makes the mechanism of pre-linking extremely
unconstrained, to the point of vacuity.

It should be observed at this point that it is not the mechanism of 'pre-linking' per se that is objectionable. As a matter of fact, the present proposal makes use of the mechanism as well. What is objectionable is using it in conjunction with regular association conventions and furthermore using it for exceptional cases.

2.3.3. The absence of positive arguments. We have arguments against representing tones on the morpheme, assuming both the unitary and the sequential analysis of contour tones. Under a unitary analysis, the theory in question misses the generalization that the association between tones and syllables is one-to-one; under a sequential analysis, on the other hand, the theory encounters insurmountable difficulties in getting the correct association between tones and tone-bearing units in contour tone cases. All these negative arguments aside, can we find any positive arguments for representing tones on the morpheme in Mandarin? The answer is no.

When we refer to the criteria that motivate the morphemic representation, we see that there is neither a restriction on tonal melodies that has to be stated on the morpheme nor is there a constant tonal melody that
stays the same in expressions of different lengths. Mandarin certainly will not be used like Mende as a language to argue for the morphemic representation of tones. Why do we then represent tones on the morpheme as well in Mandarin? The only answer seems to be that Mandarin was thought to be accommodated by the theory of morphemic representation and that the theory is a universal one.

We have found, from the discussion so far, that the facts of Mandarin are not accommodated well by the theory. Empirical difficulties ignored for the moment, I suggest that assuming a theory in the absence of negative arguments would still leave something to be desired. If we are merely satisfied that facts from a new language can be accommodated by our existing theory, we may stop looking for the interesting differences of the new language that may be better captured in some other theory. Although tones in Mandarin can be represented on the morpheme, with heavy use of exceptional pre-linking, representing them this way obliterates the differences between Mandarin and Mende. The formalism predicts that there is no difference.
2.4. Syllable as the domain of tones

2.4.1. The proposal

As I argued earlier, there is no contradiction in an autosegmental theory that represents tones on the syllable or the segment. What I am proposing then is exactly such a theory. As an alternative to representing tones on the morpheme, I put forward the following hypothesis:

Hypothesis: Tones in Mandarin are represented on the syllable/segment, on an independent tier.

Tones and syllable/segments are already linked in the lexical representation, if there are both.

Hence no association/mapping conventions are necessary.

The difference between the present theory and that of Leben (1978) can best be illustrated with the derivation of alading 'Aladdin', which has no contour tones, and putaoya 'Portugal' with all contour tones:

Leben (1978):

\[
\begin{align*}
\text{H \quad H \quad H} \\
[ \quad \text{a \quad la \quad ding} \quad ]_m \\
\text{'Aladdin'}
\end{align*}
\]

\[\text{-----} \rightarrow\]

\[
\begin{align*}
\text{H \quad H \quad H} \\
[ \quad \text{a \quad la \quad ding} \quad ]_m \\
\text{mapping applies}
\end{align*}
\]
L H L H L H  
[ ∨  ∨  ∨ ]m  
pu tao ya  
'Portugal'  

----->  
L H L H L H  
[ ∨  ∨  ∨ ]m  
pu tao ya  

Present analysis:  
H H H  
[ ]m  
la ding  

L H L H L H  
[ ∨  ∨  ∨ ]m  
pu tao ya  

The possible tonal configurations allowed in this proposal are given below:  

a) T  b) T T . .  c) T  d) T  e)  
|  \ / \ / \  |
$s s s \ldots s$  

All these configurations are no different from those allowed at some point in the derivation in the autosegmental theory of Goldsmith (1976). The (a) and (b) cases are illustrated above. Configurations like (c), although allowed by the theory, do not seem to occur underlyingly in Mandarin; but they will arise in derived environments, resulting from either the WFC or deletion and spreading rules, as shown in the following:
Therefore, if there is need to refer to a tone associated with more than one syllable in rules, it is always possible to represent them one-to-one in the lexicon and collapse them in the early part of the derivation.

The (d) and (e) configurations represent a segmentless tone and a toneless segment, respectively. They certainly exist in Mandarin, as exemplified by the reduplicative floating high tone and the aspectual marker le. Nothing in the present proposal rules them out. All the proposal says is that if there are both tones and segments/syllables, they will be linked in the lexical representation. How do these toneless segments and segmentless tones get linked, though? As there are no regular mapping rules and conventions, they are linked by language-particular rules of docking and spreading.

There are a number of advantages in adopting the present position. They are different though, depending on whether contour tones are interpreted unitarily or sequentially.

Under a unitary interpretation of contour tones, the present proposal captures the generalization that the correspondence between tones and syllables is one-to-one.
what I am proposing is a multi-tiered theory, there is no problem with having a sequence of tones on the same syllable, unlike in single-tier theories. But as the domain of tones in this theory is not a morphological unit, tones and segments are not represented separately in the lexicon. They are already linked in the lexical representation of morphemes. There is no mapping and no question of alignment. Most importantly, the 'exceptional' distribution of contour tones will no longer be a problem, nor in fact exceptional.

The present proposal is actually very similar to the position advocated by Camhi (1984), a recent dissertation on the representation of tones. The data from which Camhi draws his conclusions are from African languages, the ones that motivated the suprasegmental representation in the first place. So it is possible that the present proposal can be extended to African languages as well. But I will not claim that we should take syllable tone association as universally true. Whether languages have syllable tones or morpheme tones may be a parameter along which languages can vary.

2.4.2. Syllable tone and the OCP.

Under the present analysis, the Obligatory Contour Principle (OCP) will be systematically violated. As the syllable is the melody-bearing unit in Mandarin, there is
nothing that prevents identical tones in adjacent syllables. If the OCP is a well-motivated principle, then our analysis bears the burden of refuting the OCP as a universal principle. The question is: is OCP a well-motivated principle?

Odden (1986) convincingly shows that OCP is not a well-motivated principle. There are phonological contrasts in natural languages that violate the OCP. One such contrast, HL versus HHL, is found in the pair of words from Mende:

\[
\begin{array}{c|c|c}
\text{ngila} & \text{ngongo} & \text{ngongo} \\
\end{array}
\]

One way to solve this particular problem, as Odden pointed out, is to resort to Leben's (1978) pre-linking solution:

\[
\begin{array}{c|c|c}
\text{ngila} & \text{ngongo} & \text{ngongo} \\
\end{array}
\]

Although in this case OCP is salvaged, Odden gave other genuine counter-examples to the OCP from a number of languages and concluded that OCP should not be taken as a linguistic principle.

In the present proposal, where the syllable is the domain of tones, the OCP will be violated at the lexical
level. Since the syllable is the relevant contrastive unit, the whole paradigm of tones can be instantiated in each syllable. The contrast between ngila and ngongo is captured without problem and without the extra machinery of pre-linking:

\[
\begin{array}{ccc}
\text{HL} & \text{*HL} & \text{HHL} \\
\text{ngila} & \text{ngongo} & \text{ngongo}
\end{array}
\]

To generalize, adjacent identical tones in different syllables are preserved at the underlying level. If they should be collapsed at all, it should be achieved by language-specific rules and hence is not cost-free, as Odden showed in his paper. Therefore, the default is not to apply the OCP. In the original proposal by Leben, however, OCP produces the unmarked case. Our theory therefore is consistent with the markedness prediction of Odden's analysis of the OCP.

2.4.3. Syllable tone and the power of the theory. The present proposal, not observing the OCP, is no doubt a less restricted theory than the one I argue against. Doubts concerning the power of the theory may therefore be raised. Two things can be said in defence of the present theory.

First of all, meta-theoretical considerations cannot override the need to account for empirical facts.
In Mandarin, syllables at all positions can bear contour tones as well as level tones. Assuming the sequential interpretation of contour tones, we cannot account for the distribution of contour tones in any instantiation of the association conventions.

Furthermore, the increase in the possibility of contrasts is offset by the simpler machinery we need to posit in the present proposal. If we want to represent tones on the morpheme and account for the distribution of contour tones, we must have a more complicated theory, whether or not we assume the sequential interpretation of contour tones. Under a sequential analysis of contour tones, 'exceptional' pre-linking has to be used for unexceptional cases; under a unitary interpretation of contour tones, positing association conventions is an unnecessary complication, since the correspondence between tones and syllables is always uninterestingly one-to-one. Under the present proposal, there is no association and no pre-linking.

2.5. The Issue of Monosyllabicity.

We showed in 2.3. that polysyllabic morphemes constitute the strongest argument against representing tones on the morpheme in Mandarin. No existing tonal
association conventions can be made to work in these cases. An understandable objection may arise at this point: how strong an argument is it if it has to rely on examples that are few and far between in the language?

It is true that polysyllabic morphemes in the language are few. The examples that are representative in the native vocabulary include:

putao 'grapes', laji 'garbage', mahu 'careless'.

Examples like these are mostly foreign loan words that entered the language a long time ago. The first two examples use characters that are not used anywhere else, and the individual syllables are meaningless. In the third example, the characters denote the morphemes 'horse' and 'tiger' respectively. (Students of Chinese are always fond of the reduplicated form mamahuhu, literally 'horse-horse-tiger-tiger', with the idiomatic meaning of 'so-so'.) But since there is no regular meaning relationship between horse and tiger and being careless, we should take mahu as a simple morpheme.

Apart from these sporadic examples, there are the transliterations of foreign names, of which the following are examples:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ligen</td>
<td>mali</td>
<td>yuehan</td>
</tr>
<tr>
<td>'Reagan'</td>
<td>'Mary'</td>
<td>'John'</td>
</tr>
</tbody>
</table>
These examples are not rare, and new items are constantly being added to the language.

There has been doubt as to the monomorphemic status of these examples. Do native speakers take them to be complex or simple? It is true that in transliterating foreign names, care is taken to select characters that are either largely meaningless and reserved for transliterations or ones that are appropriate for names in Chinese. The second character used to represent mali 'Mary' for example is very often the character representing the morpheme 'beautiful' or the character used in the morpheme 'jasmine'. And reanalysis indeed happens in the case of putao 'grapes', as some people use the character for 'peach' to write the syllable tao. Despite George Kennedy's vehement attack on the 'monosyllabic myth' (1946), many people still consider Chinese as a monosyllabic language, in the sense that the syllable and the morpheme are coterminous.

Suppose then, for the sake of argument, that some polysyllabic morphemes are really compounds. Although this reduces the number of polysyllabic morphemes, the argument against representing tones on the morpheme is not significantly weakened. We have to be prepared, though, for the limiting case, where all the morphemes in
the language are analysed as monosyllabic. In that case, do we have to say that tones have to be represented on the morpheme?

It is certainly true that taking each syllable as a morpheme avoids the two problems we had earlier with representing tones on the morpheme: the one-to-one correspondence and the difficulty of mapping contour tones. In the first case, we can postulate a morpheme structure constraint to the effect that each morpheme has one tone; in the second case, the mapping of contour tones is straightforward, as long as we assume cyclic tonal association:

\[
\begin{array}{c}
L & H & L & H \\
pu & tao & ya \\
\end{array}
\]

But the fact that a theory can get out of a problem does not make it automatically attractive. The existence of polysyllabic morphemes entails, as we have shown, the impossibility of representing tones on the morpheme, given the presently available machinery. The non-existence of polysyllabic morphemes, however, does not entail the necessity of representing tones on the morpheme but is consistent with such representation. It is also consistent with representing tones on the syllable. We are by no means forced into the morphemic representation of tones by the monosyllabicity of the language.
Notes to Chapter II

1. Analyses in the autosegmental framework, such as Yip (1980), simply take the third tone to be a low tone underlyingly. The fall and the rise, especially the rise, seen in 214 are accounted for by phonetic rules applying in citation forms and pre-pausally. In 3.2, I re-examine the issue of the underlying form of the third tone and conclude that there is advantage in taking the more complex contour as the underlying form.

2. Two points are used even when a level tone is involved. No explicit claims are made about the simple/complex nature of the tones by the notation itself (Some descriptions use single digits to transcribe short level tones on checked syllables.).

3. There is a rough correlation between segmental and phonological domains and that between suprasegmental and morphological domains, when both 'segmental' and suprasegmental' are used in the domain sense. The correlation breaks down in the case of the syllable, which is suprasegmental on the one hand and phonological on the other.

4. As I will argue in chapter 5, the spreading analysis has a number of difficulties. But Yip's example is still used here to demonstrate spreading.

5. The theory that results is intermediate in tier-independence between Goldsmith's proposed theory and the 'deautosegmentalization' theory he entertains. It has less independence than Goldsmith's proposed theory because there is no mapping at the initial level and because synchronization between tiers already exists at the level of lexical representation. It has more independence than the 'deautosegmentalization theory' because the tiers are independent in the lexical representation and no 'tier-separation' is needed.
6. My position on the status of contour tones is similar to that advanced by Yip (1988), which holds that while contour features like [+rising], [+contour] are unnecessary, contour tones have to be treated at some level of the grammar as unitary. More specifically, contour tones should be analysed sequentially at the post-lexical level, as evidenced by Mandarin Second-to-First Tone Rule (LH---\rightarrow HH/[H ]g), which has to refer to the level components of the contours; but at the lexical level, contours must be regarded as wholes. In the mapping between tones and syllable/segments, with which the current argument is concerned, contour tones behave unitarily.
CHAPTER III

TONE SANDHI IN MANDARIN

After discussion of the unit of representation in Mandarin, let us now turn to tone sandhi rules in the language. The tone sandhi phenomena in Mandarin Chinese are quite well known now in the phonological literature, thanks in large part to the first generative phonology of the language by C. Cheng (1973), which makes the facts accessible to phonologists in a format that is familiar to them. In particular, the so-called Third Tone Rule, by which a third tone changes to a second tone when another third tone follows, has featured prominently in theoretical discussions. There are two main reasons for the importance of the Third Tone Rule: there are problems with the rule's mode of application, and the rule has complex morphosyntactic and prosodic conditions on its application. Discussion of the former issue can be found, for example, in such discussions on rule application as Howard (1972) and Anderson (1974). The rule's role in the syntax-phonology interface has been discussed in two recent contributions to the subject: Kaisse (1985) and Selkirk (1984). The prosodic domain of the rule's
application has been explored by Chen (1986), Shih (1986) and L. Cheng (1987). An autosegmental analysis of the main Mandarin tone sandhi rules can be found in Yip (1980).

Despite the appearance of its being overworked, the topic is nowhere near exhausted. Most analyses in the generative framework have centered on the Third Tone Rule. While the present chapter includes a discussion of the Third Tone Rule as well, it also presents other less known tonal changes. One of these tonal changes, the Second-to-First Tone Rule, actually turns out to be very important in revealing the prosodic structure of the language.

The present chapter is a detailed analysis of the various tonal rules in Mandarin. The aim is to show that prosodic structures, be they metrical structures or prosodic domains, play an important part in the tonal rules of Mandarin. Reanalysis of a number of rules previously formulated with tonal environments shows that either they are conditioned directly by metrical relations or their domains of application are constructed by reference to strength relationships.

Another important point, to be learned only by looking at the whole range of tone sandhi facts, is that what have been regarded and expressed as the same kind of tonal rules, i.e. tone sandhi, turn out to be very different from each other.

The chapter is organized in the following way. The
first section deals with a class of tonal alternations that should not be taken as tone sandhi in the strict sense. In this class, we discuss the tonal changes that are found in specific morphemes, yi 'one', qi 'seven', ba 'eight', and bu 'not'; we will then look at the tonal change in the morphological process of reduplication.

The second section discusses the half tone rules. I argue against previous formulations with tonal conditioning factors and suggest that these tonal changes are induced by prosodic prominence.

The third section similarly argues against previous formulations of the Second-to-First Tone Rule that use only tonal information in the structural description. Yip's (1980) formulation with metrical information is also found to be unsatisfactory. A reformulation making use of prosodic domain is given. It will be seen that the prosodic domain motivated on the basis of this rule will also correspond to the domain in the Mandarin Third Tone Rule.

The last section reviews the previous analyses of the Third Tone Rule. It will be concluded that the prosodic domain is necessary for the application of the rule. But the specific analysis of Shih (1986) is unsatisfactory in a number of ways. A modular approach to the problem, with separation of lexical from post-lexical application and the introduction of pragmatically
controlled phonological phrasing, will be suggested.

3.1. Morphological tonal changes

3.1.1. yi, qi, ba, bu. Apart from the general tone sandhi changes with which most of the chapter is concerned, there are a few tonal changes that are applicable to only a few morphemes, or even one. The morpheme yi 'one', which has the first tone in isolation and prepausally, has the fourth tone when the following syllable is not a fourth tone. When the following syllable is in the fourth tone, the morpheme has the second tone. The morpheme bu 'not' has the second tone when followed by a fourth tone but has the fourth tone elsewhere. For some speakers, the morphemes qi 'seven' and ba 'eight', which otherwise have the first tone, will have the second tone in front of the fourth tone. We can represent the distribution of the morphemes as follows:

<table>
<thead>
<tr>
<th></th>
<th>55</th>
<th>51</th>
<th>55</th>
<th>55</th>
</tr>
</thead>
<tbody>
<tr>
<td>yi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'one'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bu</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'not'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qi</td>
<td></td>
<td>51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'seven'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ba</td>
<td></td>
<td></td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>'eight'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>55</th>
<th>51</th>
<th>55</th>
</tr>
</thead>
<tbody>
<tr>
<td>di.yi</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'first'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wo.bu</td>
<td>55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'I won't'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shi.qi</td>
<td></td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>'17'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shi.ba</td>
<td></td>
<td></td>
<td>55</td>
</tr>
<tr>
<td>'18'</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>51</th>
<th>55</th>
<th>55</th>
<th>55</th>
<th>55</th>
</tr>
</thead>
<tbody>
<tr>
<td>yi.zhang</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'one piece'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bu.zhi</td>
<td>55</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'not know'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qi.tian</td>
<td>55</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'7 days'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ba.tian</td>
<td></td>
<td>55</td>
<td>55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'8 days'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>51</th>
<th>35</th>
<th>51</th>
<th>35</th>
<th>55</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>yi.pan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'one plate'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bu.shi</td>
<td>55</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'not know'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qi.ping</td>
<td></td>
<td>55</td>
<td>35</td>
<td>55</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>'7 bottles'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ba.ping</td>
<td></td>
<td></td>
<td>55</td>
<td>35</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>'8 bottles'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
It is generally agreed that only the morphemes yi, bu, qi and ba are involved in these alternations. But I find that for some speakers the pronunciation of the morphemes san 'three' patterns with that of qi and ba, though this has never been reported in the literature as far as I know. It is interesting now to note that with the addition of san a natural class is formed: yi, qi, ba, and san are the only numerals below ten that have the first tone.

In discussing the variants of the tone in yi, Jin (1979) points to the variation as being governed by specific constructions. The variants are found in certain number constructions, among them noun phrases involving numbers and classifiers (but not ordinal numbers) and adverbial expressions with numbers and frequency markers. Even in the same tonal environments, we find cases of both alternation, as in the first example below, and lack of alternation, as in the second.

51 55
yi tian yi fen wei er
'one day' one divides into two

51 55 35 51
yi fen wei er
one divides into two
'look at two sides of things'
Of these morphemes, only yi has three variants. Its citation/pre-pausal variant is different from both the variant occurring in front of the first, second and the third tone and the variant that occurs only in front of the fourth tone. We may take the citation/pre-pausal variant to be derived from the variant that occurs in front of the other three tones, i.e. 51 → 55/___#.

The other morphemes have two variants, one in front of the fourth tone, 35, and one elsewhere, 55 or 51 depending on the morpheme. We can therefore posit a lexical rule\(^2\) that changes a non-rising tone to a rising tone in front of a falling tone:

\[
\begin{align*}
55 & \rightarrow 35 / \_\_\_ 51 \\
51 & \rightarrow 55 / \_\_\_ 51
\end{align*}
\]

This is not a general rule, as liu 'six' (with the tone 51) does not change to 35 before dui 'measure word', another 35-toned morpheme. Nor do er 'two' and si 'four', both with the 51 tone.

The change from 51 to 35 in front of another 51 seems to be a more general rule. It was recently reported by Lin (1985). Following are some examples from Lin:

\begin{verbatim}
51 51 ----> 35 51
dian shi dian shi 'television'
51 51 ----> 35 51
xian zai xian zai 'now'
\end{verbatim}
According to Lin, the 35 may reflect the original sandhi form or the original state of the fourth tone. This phenomenon also is variable according to particular lexical items, the sociolinguistic affinity of the speakers and so on. He denies any dialectal interference, however. While some native speakers from Beijing I have consulted claim that they are not aware of the tonal change, the phenomenon would not be surprising, since it already exists to a limited extent in the tonal changes of yi, qi, ba, bu, and possibly san.

Lin's postulation of the rule was challenged by He (1988), who explained the change from 51 to 35 as due to intonational effects. The arguments He forwarded are the following: the change does not occur in reading of word lists but in sentences; the change is not lexically specific; the change does not occur on stressed syllables.

3.1.2. **H tone in Reduplication.** There is a tonal change that is closely associated with reduplication. Monosyllabic adjectives or adverbs can be reduplicated to give an 'intensifying effect'. Sometimes, the second syllable in the reduplicated form takes on the invariable
first tone, regardless of the original tone of the syllable. When this happens, there is lengthening of the syllable and usually r-suffixation as well. The overall effect of H (55) tone plus suffixation is that of 'vivifying'. Below are two examples:

\[
\begin{align*}
51 & 51 & 55 \\
\text{man} & \text{-------> man} & \text{manr} \\
& \text{'}slow' \\
214 & 214 & 55 \\
\text{xiao} & \text{-------> xiao} & \text{xiaor} \\
& \text{'}small' \\
\end{align*}
\]

This tonal change should not be called tone sandhi, since there is no tonal environment responsible for the change. Instead, the change in tone is conditioned by the grammatical process, be it word-formation or a syntactic construction. The change should best be described as the affixation of a morpheme, consisting entirely of a H toneme and with the meaning of 'vivification'.

The floating H tone has been used by Yip (1980) as an argument for the autosegmental nature of tones. Yip also argues that the change is best described by reduplicating the segments only and docking the floating tone onto the reduplicated syllable, without having to delete the tone on the reduplicated syllable. This she uses as further confirmation of the independence of the tonal tier.

This analysis, though simple, is problematic. The
invariable H tone attachment is only optional and depends largely on the semantics of particular lexical items (Xu 1978). The H tone is responsible for the vivifying effect and if the semantics of the lexical item is incompatible with such an effect, the H tone will not be attached. The following examples cannot have the H tone:

\[
\begin{align*}
&51 \quad 51 \quad 51 \\
&lu \quad -------> \quad lu \quad lu \\
&'green' \\
&51 \quad 51 \quad 51 \\
&mo \quad -------> \quad mo \quad mo \\
&'silently'
\end{align*}
\]

To be able to get these forms, the tones have to be reduplicated together with the segments, with the tones on the second syllable then deleted when a H tone is attached. The rules of reduplication and attachment of the H tone have to be separate to predict the optional attachment of the H tone as well as the semantic difference.

If we want to accommodate both the cases with H tone and those without in Yip's analysis, we would have to set up two separate reduplication processes. Suppose two separate reduplication operations are set up and the H tone attachment is optional, then we will predict not only second syllables with original tones and the invariable H tone but also second syllables without tones at all. Such forms have not been observed in the language. To save the
analysis, we would have to ensure that when toneless reduplication occurs, H tone attachment has to occur too. Such an analysis is not any simpler than one that posits deletion.

The present analysis is consistent with what we concluded in the last chapter, that is, Mandarin is a syllable tone language. In a syllable tone language, the reduplication of the whole syllable, tones and segments in all, is the unmarked situation. In Yip's analysis, to accommodate whole syllable reduplication, tonal association has to occur before reduplication.

3.2. Purely metrically conditioned tonal rules

There are two tone rules in Mandarin Chinese that can be shown to be directly accounted for by the metrical patterns of Mandarin Chinese. That is to say, the conditioning factor in the rules should not be tones, as is traditionally assumed, but the stress pattern. However, before these tonal rules can be discussed, an exposition of the stress system of Mandarin is in order.

In addition to tones, Mandarin also a stress system. Disyllabic compounds in Chinese are divided into two kinds, depending on where the stress falls. Forms having final stress are called iambic forms and those having initial stress are called trochaic forms. Some examples
of iambic forms and trochaic forms are given below (digits indicate the degree of stress):

<table>
<thead>
<tr>
<th>IAMBIC</th>
<th>TROCHAIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 1 xue xi</td>
<td>1 0 jie jie</td>
</tr>
<tr>
<td>'study'</td>
<td>'sister'</td>
</tr>
<tr>
<td>2 1 gao xing</td>
<td>1 0 xi huan</td>
</tr>
<tr>
<td>'happy'</td>
<td>'like'</td>
</tr>
</tbody>
</table>

Trisyllabic forms with full-toned final syllable have the primary stress on the final syllable, secondary stress on the first syllable and the weakest stress on the second syllable.

| 2 3 1 san nian ji |
| third year grade  |
| 'the third grade' |

| 2 3 1 cong you bing |
| onion oil cake     |
| 'a kind of pancake' |

At the phrasal level, the stress pattern is right-dominant.

Yip (1980) uses the metrical framework of Liberman and Prince (1976) to describe the stress pattern in Mandarin Chinese. According to her system, there are only two levels in the stress tree, the level of the foot and the phrasal level. There are monosyllabic feet, which are constructed on a full tone, and polysyllabic feet, in
which the initial syllable is strong, followed by one or more toneless syllables. Such polysyllabic feet represent the trochaic forms. Feet are grouped together into phrases. And at the phrasal level, the stress pattern is right-dominant. The iambic forms, both compounds and phrases, are represented at the phrasal level of the metrical tree. An example each of a trochaic form (a), an iambic form (b), and a trisyllabic form with a final toned syllable (c) is given below:

(a)  
\[
\begin{array}{c}
\text{jie jie} \\
\text{'sister'}
\end{array}
\]

(b)  
\[
\begin{array}{c}
\text{f f f} \\
\text{f f f}
\end{array}
\]

(c)  
\[
\begin{array}{c}
\text{xue xi san nian ji} \\
\text{'study' 'the third grade'}
\end{array}
\]

In the following example given by Yip, only the phrase level is involved (p.190):

\[
\begin{array}{c}
\text{lao li mai hao jiu} \\
\text{old Li buy good wine}
\end{array}
\]

In the following tree, both the foot level and the phrase level are involved:
Later on, I will argue that Yip's system is too restrictive in allowing only two levels.

3.2.1. The Half Third Tone Rule. The first half tone rule is the so-called Half Third Tone Rule, which C. Cheng (1973:43) states as follows:

Third tone becomes low-falling when preceding any tone other than third tone.

We find similar renditions of the rule in Chao (1968) and Li and Thompson (1981). The low-falling contour that results from the application of the rule is called 'half-third tone' by Chao (1948). We can rewrite the rule in the following way:

\[
\begin{array}{c}
\text{214} \\
\text{21} \\
\text{hao xin hao ren hao shi} \\
\text{'good heart' 'good person' 'good things'}
\end{array}
\]

Examples showing the effect of the rule are given below:

\[
\begin{array}{cccc}
214 & 55 & 214 & 35 \\
21 & 55 & 21 & 35 \\
\text{hao xin} & \text{hao ren} & \text{hao shi} \\
\text{'good heart'} & \text{'good person'} & \text{'good things'}
\end{array}
\]
Most discussions of the rule explain it as the third tone losing its rising part of the contour when anything other than the third tone follows.

The rule as it is written is quite unrevealing as to the nature of the change. The conditioning factor in this formulation is definitely strange. It is the complement of the third tone. When the third tone follows another third tone, the Third Tone Rule applies, changing the first third tone to a second tone. The effect of the disjunctive environment then is to exclude the third tone. This means that 55/35/51 may not be the real conditioning factor. Even if it were, in what way does it condition the change? Is the rule assimilation? Or is it dissimilation? It is hard to see how the complement of the third tone forms a natural class.

Suppose we said that the rule is unconditional and state the rule in the following way:

2 1 4 ----> 2 1

If we order the Third Tone Rule before this rule, whatever instances of the third tone not changed by the Third Tone Rule will undergo this rule. This cannot be right either, because the rule must not apply in certain contexts, at least pre-pausally and in citation forms. Therefore, although the conditioning factor for the rule is not
tonal, it is a conditioned rule.

What is the condition then? The rule affects the tone of the first syllable in a disyllabic sequence. The tones on the second syllable of such sequences are left unchanged. Now, we know that the majority of disyllabic compounds and phrases are iambic forms and the stress pattern of disyllabic iambic forms is [w s]. Thus, the real condition is the stress pattern; the rule is a weakening rule, contingent on weak stress. This makes phonetic sense, because the loss of a part of the complete contour is presumably what weakening can do. We can state the Half Third Tone Rule in a more sensible manner this way:

```
/ \    \\
  w s
```

214 ----> 21 /

The rule applies after the Third Tone Rule, which changes 214 to 35 when another 214 follows. Since the Third Tone Rule applies both lexically and post-lexically, the metrically conditioned weakening rule is a post-lexical process.

3.2.2. The underlying form of the third tone. So far, the discussion has assumed that the full contour is the underlying tone and the half contour is the result of a rule. This assumption is not held by everybody, however.
Yip (1980), for example takes LL (the equivalent of 21) as the underlying form and derives LLH (214) by a rule which inserts the H when the syllable is immediately dominated by 's'. The direction of change then is exactly the opposite of what we have been assuming.

While the difference in the direction of change is independently interesting, both Yip's rule and the one given here describe the same relationship between stress and tone change. It is interesting to pursue the question of directionality a bit further, however.

The analysis of the third tone as being simply low underlyingly is based on the limited distribution of the full contour of the third tone. If it is found that the half third tone actually occurs in most environments, it is then reasonable to take it as the basic variant. The question is how limited the distribution of the full contour is. There is no question that it occurs in citation form and pre-pausally. If it can occur only in this environment, then its distribution is indeed limited. But empirical confirmation is needed that a full contour, albeit not quite the full citation contour, does not occur anywhere else.

Yip's analysis of the third tone is consistent with her metrical analysis of Mandarin, according to which the only syllables immediately dominated by 's' in the sentence are those in absolute final position and those
before a neutral tone in trochaic forms. The 's' positions in trochaic forms are underlined in the tree below:

\[
\text{phrase: } \quad / \quad \backslash \\
\quad W \quad s \\
\text{foot: } \quad / \quad \backslash \\
\quad s \quad W \quad s \quad W \\
\]

Since trochaic forms are in the minority in Mandarin lexical items, the context for the occurrence of the full contour of the third tone will be very limited there. Following Halle and Vergnaud (1978), Yip assumes that within a single level, trees cannot branch in two different directions. Since the stress pattern is right-dominant at the phrase level and left-dominant at the foot level, the phrasal tree will be right-branching all the way but the foot tree will be left-branching all the way. It follows then that the only foot immediately dominated by 's' in the phrasal tree is at the end of the sentence. As a matter of fact, we can use one of Yip's examples (190) to illustrate this:

\[
/ \quad \backslash \\
\quad s \\
\quad / \quad \backslash \\
\quad s \\
\quad / \quad \backslash \\
\quad s \\
\quad W \quad W \quad W \quad W \quad s \\
lao \quad li \quad mai \quad hao \quad jiu \\
old \quad Li \quad buy \quad good \quad wine \\
2 \quad 3 \quad 4 \quad 5 \quad 1
\]
The stress values given above are arrived at according to the algorithm of Liberman & Prince (1976). The underlined positions in the trees are the only places where the full contour of the third tone will surface.

It seems that Yip's system is too restrictive. In the tree above, the stress values are [2 3 4 5 1], where 1 is strongest and 5 weakest. The decreasing of syllable strengths up to the penultimate syllable cannot accommodate iambic disyllabic forms at any position other than the final two syllables, as iambic forms have the [w s] pattern. lao li 'old Li', an iambic form, cannot be pronounced with the first syllable stronger than the second syllable, as predicted by Yip's system. In order to maintain the constraint against branching weak nodes within a single level, I suggest that we introduce an additional higher level in Mandarin metrical trees. The following sentence illustrates the revised system:

```
phrase:         / \  \\
               w  \\          \\
--- s
?::           / \  /\  \\
            w s w s
-------------
foot: lao Li mai yao
       4 3 2 1
'Old li buy medicine'
```

The numbers in the tree indicate relative stress values. In this tree, lao is weaker than Li, which is much closer
to facts than the opposite ranking in a Yip's system:

phrase:

foot: lao Li mai yao
      2 3 4 1
      'Old li buy medicine'

The digression on metrical trees tries to show that with a different, probably more realistic, metrical tree, the strong position is no longer restricted to two places (the absolute final position and before neutral tones in trochaic forms). If that is the case, the full contour of the third tone may not be that limited after all, and taking it as the underlying form is then a viable solution. In fact, if the revised metrical tree is correct, it would be hard to make sense of the Half-Third Tone Rule if it does not consistently have the fuller contour of the third tone at all strong positions, not just two of them.

3.2.3. The Half Fourth Tone Rule.

The second rule that behaves like the Half-Third Tone Rule was originally proposed by Chao (1968), but it has not been very widely discussed and hence does not have
a name. I will call it the Half Fourth Tone Rule.
Translating from Chao's prose description, I will state
the rule as follows:

\[ 51 \longrightarrow 53 / \quad 51 \]

Examples showing the effect of this rule are given below:

\[
\begin{array}{cccccc}
51 & 51 & 51 & 51 & 51 & 51 \\
\longrightarrow & 53 & 51 & 53 & 51 & 53 & 51 \\
da & shu & fu & jin & mu & di \\
& 'big tree' & 'nearby' & 'cemetery'
\end{array}
\]

Although it might be suggested that this is a
dissimilation rule, I will say that it is simply weakening
due to the weaker stress of the first syllable and will
state the rule as:

\[
\begin{array}{c}
51 \longrightarrow 53 / \\
\wedge \quad s
\end{array}
\]

It may be objected that in the original formulation of the
rule, the tonal environment, i.e., the fourth tone, is not
irrelevant, as the new reformulation suggests. But I
think the reason that we have 51 as the condition in the
old formulation is due to the fact that it is easier to
identify the changed tone when an identical sequence is
involved.

I hypothesize that the weakening process is a
general one applying to all tones at the weak position. We can then state the rule thus:

\[
\begin{align*}
\text{T} & \quad \rightarrow \quad T_w / \quad \text{(T}_w=\text{weakened tone)}
\end{align*}
\]

3.3. The Second-to-First tonal rule.

There is a sandhi rule in Mandarin, operative in conversational speech, which is stated by Cheng (1973:44) as:

"In fast conversational speed, a second tone becomes first when preceded by first or second tone and followed by any tone other than the neutral tone"

Hyman (1975) formulated the rule in the following way:

\[
\begin{align*}
35 & \quad \rightarrow \quad 55 / \begin{cases} 35 \\ 55 \end{cases} \quad \text{T}
\end{align*}
\]

I now provide examples showing the effect of this tonal rule (Chao 1968).

First, examples with the first syllable bearing the first tone (55):

\[
\begin{align*}
55 & \quad 35 \quad 55 \\
\rightarrow & \quad 55 \quad 55 \quad 55 \\
\text{xi yang sen} & \quad \text{western ginseng}
\end{align*}
\]
Examples with the first syllable bearing the second tone (35):

------>  35  35  55
shei neng fei
who can fly

------>  35  35  35
hai mei wan
still not finished

------>  35  35  214
you zha quei
oil fried dough

------>  35  35  51
long fu si
Long Fu temple

There are three other examples Chao gave; with these added, all the possibilities are exhausted. The first two show that the syllable undergoing the Second-to-First Tone Rule is derived from a third tone by the Third Tone Rule:
The third one shows that both the syllable undergoing the change and the conditioning factor to the left are derived from the third tone by the Third Tone Rule:

3.3.1. Problems with the statement of the rule.

Let us examine Hyman's formulation more closely:

\[
35 \quad -----\rightarrow 55 / \begin{cases} 35 \\ 55 \end{cases} \quad T
\]

The rule is explained by Hyman as a case of assimilation. The change can be seen as assimilation easily, according to Hyman, if we write the rule out in terms of pitch levels rather than with contour tones features. In the input, \([35 \; 35 \; T]\), for example, the first tone of the second syllable, with the pitch level of 3, is
wedged between the 5 of the preceding syllable and the 5 immediately following and hence becomes 5 as well. The same is true in the case of [55 35 T]. Thus Hyman's statement brings out the assimilatory nature of the change and is an improvement over Cheng's statement of the rule in terms of tone types.

There are a number of other curious things we can observe about this rule.

First of all, it is not apparent what the T is doing in the rule, since the assimilation is between the first two syllables. At the same time the presence of T is crucial, because the assimilation does not occur in a disyllabic sequence, however fast we pronounce the following examples:

\[
\begin{array}{c}
\text{hua ping} \\
\text{flower vase} \\
\text{55} \ 35
\end{array}
\Rightarrow
\begin{array}{c}
\text{yuan hou} \\
\text{ape monkey} \\
\text{35} \ 35
\end{array}
\]

Since the presence of the third tone is necessary, we expect that it is part of the conditioning factor. In the formulations of the rule seen so far, this part of the conditioning factor is stated tonally as well, being any tone other than the neutral tone. Since a third syllable can have the second tone as well, we expect that assimilation can occur in the sequence of T 5 35. But
this is not the case, as shown below:

\[
\begin{align*}
*35 & \rightarrow 55/35 \{35\} \\
 35 & 35 35 \\
------ & \rightarrow *35 35 55 \\
35 & 55 35 \\
------ & \rightarrow *35 55 55 \\
hai & mei wan \\
still & not finish \\
------ & \rightarrow 55 35 35 \\
------ & \rightarrow *55 35 55 \\
55 & 55 35 \\
------ & \rightarrow *55 55 55 \\
san & nian ji \\
third grade
\end{align*}
\]

In the examples above, the second syllables are changed to the first tone but the third syllables are not.

This Hyman did not explain. Nor did he explain why there should be the qualification in the rule that the third syllable should be in any tone other than the neutral one.

The rule as stated is then a paradigm example of a formulation that seems to cover the facts but does not reveal the nature of the process involved.

3.3.2. **Stress patterns**

We begin to get a clue of the real contribution of the third syllable when we consider that the equivalent of saying 'any tone but the neutral tone' is 'stressed syllable'. In Mandarin, the neutral tone is associated
with stressless syllables and full-toned syllables are all stressed. The conditioning factor in the third syllable is not tonal at all. Now let us consider some facts about the stress pattern of trisyllabic words/phrases.

Regardless of morphosyntactic structure, in trisyllabic units with non-neutral third syllables, as described by Chao (1968), the stress pattern is:

secondary tertiary primary (or: 2 3 1)

\[
\begin{array}{c}
\$ \\
\$ \\
\$ \\
\end{array}
\]

The second syllable is the shortest and the third syllable is the longest. In terms of metrical trees, the above stress pattern can be represented as:

```
/ \  \\
/   s  \\
/    / \ \\
W  W  s  \\
2  3  1
```

An alternative structure, while giving the desired surface pattern, violates the restriction in metrical theory that weak nodes do not branch:

```
/ \  \\
/ \  \\
/   s  \\
/   / \\
W  W  s  \\
2  3  1
```
When the third syllable is a neutral-tone syllable, the stress pattern becomes:

secondary primary stressless (or 2 1 0)

\[
\begin{array}{c}
\text{s} \\
\text{s} \\
\text{s}
\end{array}
\]

\[
/ \big/ \\
/ \big/ \\
/ \big/ \\
\text{s} \text{s} \text{s}
\]

\[
2 \ 1 \ 0
\]

3.3.3. Yip's statement of the rule.

In Yip (1980:188), we find the Second-to-First Tone Rule stated as follows:

\[
\begin{array}{c}
\text{w} \\
\text{L} \rightarrow \text{H} / \text{H} \_ \text{H}
\end{array}
\]

Compare this formulation of the rule with Hyman's formulation:

\[
35 \rightarrow 55 \left\{ \begin{array}{c} 55 \\ 35 \end{array} \right\} \text{T}
\]

For ease of comparison, Hyman's rule can be rewritten as:

\[
\begin{array}{c}
\text{L} \rightarrow \text{H} / \text{H} \_ \text{H} \text{T}
\end{array}
\]

One immediate difference between Yip's rule and Hyman's rule is the absence of the T in Yip's rule. This feature of Yip's analysis captures the insight that the tone of the third syllable is not really relevant.
Instead, the conditioning factor is taken to be the stress of the changed syllable, which is \( w \).

While Yip's formulation is correct in that the tone on the third syllable is not relevant to the rule, her rule also has a number of problems.

Yip uses the metrical weak node as part of the conditioning factor, but since the strong-weak relation is a relative matter, we do not know what the weak syllable is weak in relation to. Is it to the following syllable, as illustrated below?

\[
\begin{array}{c}
/ \ \ \\
/ \\
w \wrightarrow s \\
2 \ 3 \ 1
\end{array}
\]

Or is the metrically weak syllable weak relative to the preceding syllable (as we can see, both trees have the stress values of \([2 \ 3 \ 1]\))?  

\[
\begin{array}{c}
/ \ \\
/ \ \\
/ \ \\
w \\
s \leftarrow w \ \ s \\
2 \ 3 \ 1
\end{array}
\]

The answer seems to be that although the relation of the weak node to other nodes is not specified in the rule, it can be predicted from the other facts in the language. As a matter of fact, Yip claims that the right-branching tree above is the one for these trisyllabic forms in the language, regardless of syntactic structure.
If in fact the right-branching tree is the correct one, then there is something strange about the rule. Let's examine how the structural description of the rule is met:

```
  / \
 /   \
W W s
/ \ / \ 
H L H
--->
HHH
```

In the above tree, the structural description is met and the L will change to H. In a sequence of three syllables A, B and C, where the tonal part of the structural description is met both between A and B and between B and C, however, there is no change between B and C.

```
  / \
 /   \
W W s
/ \ / \ 
H L H L H
--->
HHH L H
--->
*HHHH H
```

The strange thing is that the tone rule does not apply within a metrical constituent (the underlined s) but between such constituents (the two underlined nodes). And the rule makes the wrong predictions. Let us look at the following example, with the tree drawn according to Yip's system: 
The underlined syllable meets the structural description of Yip's rule, but as we can see it does not change to 55.

Suppose we take the second interpretation of Yip's rule, i.e. with the w node in her rule construed with the preceding syllable.

We can then explain why sandhi is not possible between the second and the third syllable, even when the tonal environment is present:

Sandhi does not occur between the last two syllables because they do not form a constituent. Similarly, if the tree for Lao Jin chan jiu 'Old Jin craves wine' is as
follows, we can also explain why sandhi in the sentence is impossible:

```
/ \       
/ \      
w s      w s
/ \      / \ 
 w s w s  w s
lao jin chan jiu
old Jin crave wine
214 55 35 214
--->* 214 55 55 214
```

As the third syllable does not form a metrical constituent with the second syllable, it therefore does not change.

Although the second interpretation of Yip's rule makes more sense, the metrical trees that allow such interpretation are certainly not what Yip's system licenses. They all have branching weak nodes. Since Yip assumes Halle and Vergnaud's (1978) restriction against bi-directional branching within a single level, these trees would be ruled out.

But as I argue earlier, Yip's system with only two levels, the level of foot and the level of phrase, is too restricted. It cannot accommodate any iambic forms that are not in the final position, since a consistently right-branching tree predicted by Yip's system gives rise to a descending scale of stress values up to the final strong syllable, but iambic forms have the stress pattern of [w s]. According to Yip's system, the stress pattern for Lao Jin chan jin 'Old Jin craves wine' is as follows:
This pronunciation is very unlikely. Our version of the tree seems to give the right stress pattern:

```
    / \  \
   /   \ \
  /     \ \
 w     w  \
```

lao jin chan jiu
old Jin crave wine
2 3 4 1

Our solution is to introduce the word level in Mandarin metrical trees. This move still retains the constraint against branching weak nodes within a single level.

Independent of the introduction of the word level, left-branching metrical trees can also arise through restructuring in fast speech. The Second-to-First Tone Rule is a fast speech rule; and there is no reason why the metrical structure should be the same under all conditions.

The restructuring process is given below:

```
    / \  \
   /   \ \
  /     \ \
 w     w  \
```
The output might not violate the spirit of metrical phonology after all, if in fast speech the first two syllables are grouped into one foot. There will then be no branching weak nodes within a single level. Properly restricted to fast speech, the mechanism of restructuring need not entail too much power.

The reinterpretation of Yip's formulation predicts that the tonal assimilation rule in question will occur in metrical constituents. While this seems more satisfying than applying the rule only between metrical constituents, there arises the question of why a fast speech rule applies in metrical constituents. Why not apply the rule in a prosodic domain such as the foot and the phonological word in the sense of Hayes (1984) and Nespor and Vogel (1986)? If we can show that there are other rules that refer to the same domain as that in which the Second-to-First Tone Rule applies, we have a principled reason to state the rule in terms of prosodic domains rather than metrically. There is in fact such evidence.

3.3.4. Prosodic Domains

3.3.4.1. Lenition and resyllabification.

Unstressed syllables can undergo segmental lenition (Lin, 1963). In the extreme case, resyllabification
occurs, as seen in the two examples below:

dou fu ----> douf
bean curd

xin si ----> xins
heart thought

In fast speech, we find lenition and resyllabification in the second syllable of trisyllabic sequences, the same ones that exhibit tonal assimilation. When resyllabification occurs, what is left of the second syllable joins the first syllable, rather than the third syllable. The following examples are from Lin (525-527):

a. bai shi qiao
   ----> bai rζ qiao
   white stone bridge

b. tu mo xingr
   ----> tum xingr
   'spit'

The following example is mine:

jia na da
   ----> jian . da
   'Canada'

In all these examples, the second syllable leans toward the first syllable and the first two syllables become a prosodic unit of some kind:

[ [ $ $ ] $ ]
That the prosodic grouping of the first and the second syllable is not due to morphosyntactic structure is obvious. Jia na da is a transliteration of 'Canada' and there is no morphosyntactic structure to speak of at all.

When the third syllable is the neutral tone, however, lenition of the second syllable is no longer possible:

\[
\begin{array}{ccc}
35 & 35 & N \\
bai & shi & zir \\
\text{---*} & bai & ri \\
\text{white stone} & \text{affix}
\end{array}
\]

The toneless third syllable leans on the second syllable, and the prosodic organization is:

\[
[\$ [\$ \$ ] ]
\]

It is now quite easy to understand why when the third syllable has the non-neutral tone, tonal assimilation applies: it is because the first two syllables are brought into the same prosodic unit. When the third syllable has the neutral tone, however, the first and second syllable no longer belong to the same prosodic unit. This explains why the assimilation rule does not apply then. And since the third syllable has the neutral tone, assimilation will not occur between the second and the third syllable even if they are in the same prosodic unit.
Our prosodic explanation can also explain why a disyllabic sequence does not undergo the rule, as disyllabic forms have the stress pattern of:

\[
\text{secondary primary} \\
\$ \quad \$ 
\]

Neither syllable is particularly weak and leans against the other. Each forms a prosodic unit of its own.

\[
[ \text{[\$]} \quad [\text{\$}] \text{ ]}
\]

And lenition of the second syllable in disyllabic forms is also impossible, if this syllable is not already toneless underlyingly or if the syllable has not lost its tone by a lexical tone deletion rule:

\[
------> \quad \text{bai} \quad \text{shi} \\
\text{* bai} \quad \text{ri} \\
\text{white} \quad \text{stone}
\]

We have seen that the tonal assimilation rule and the lenition rule apply in similar conditions. Our explanation then is that in fast speech prosodic unit reformation occurs, which brings about both lenition and assimilation.

3.3.4.2. Domains for the Third Tone Rule. Another
argument for having the following prosodic organization below for trisyllabic forms with final toned syllables comes from domains for the Third Tone Rule:

$ [ \ [ \$\ ]\ $ ]$

To anticipate discussion of the domains for the Third Tone Rule, in trisyllabic forms with no internal structure the first two syllables are grouped together by a prosodic domain formation rule for the purpose of applying the Third Tone Rule correctly. The rule scans from the beginning of the sequence and groups together disyllables into tonal domains.

Thus, the domains for the Third Tone Rule and those for the First-to-Second Tone Rule reveal the same underlying rhythmic structures of the language, which favor the grouping of the first two syllables in a trisyllabic sequence. The difference between the two rules is that in the Second-to-First Tone Rule, the domain formation does not respect morphosyntactic structure; but in the normal application of the Third Tone Rule, domain formation does respect immediate constituent structure. This difference is explained by the fact that the Second-to-First Tone Rule is a fast speech rule while the Third Tone Rule can be a lexical rule as well. In fact, I will argue that in fast speech application of the Third Tone
Rule, immediate constituent structure is again ignored and the first two syllables always form a unit.

3.3.5. Restatement

With explicit reference to the prosodic domain, we will then state the Second-to-First Tone Rule in the following way:

\[ L \rightarrow H / [ H \underline{\text{__}} H ] \text{foot}^7 \]

3.4. The Third Tone Rule.

A statement of the Third Tone Rule is given below:

\[ 3 \longrightarrow 2 / \underline{\text{_____}} 3 \]

where 3 and 2 denote the third tone and the second tone respectively. The third tone is a dipping tone\(^8\) with initial fall followed by a rise, whose contour is described as 214 in the Chinese tonal literature. The second tone is a rising tone, having the contour of 35. The first tone and the fourth tone, with the contours of 55 and 51, will be represented respectively as 1 and 4.

3.4.1. The problem.

The application of the rule in disyllabic sequences
is relatively unproblematic. This is equally true in the lexical and post-lexical application of the rule. In compounds and phrases of longer sequences, however, the application of the rule is no longer straightforward. Cheng (1973) found that in a sentence comprising entirely third-tone syllables like the one below, different pronunciations are possible, depending on the rate of speech and other 'attitudinal factors':

```
lao li mai hao jiu
Old Li buy good wine
```

```
3 3 3 3 3

-------->

a. 2 3 3 2 3
b. 2 2 3 2 3
c. 2 3 2 2 3
d. 2 2 2 2 3
```

The data are problematic for any simple scheme tinkering with only the mode of application. A free reapplication of the rule will not work. It predicts that only sequences without two 3s in a row will be generated. This rules out (a), where a sequence of 3s remains. And simultaneous application of the rule to the whole sentence will give (d) only. The same problem exists for iterative application from left-to-right. Obligatory, cyclic application on the morphosyntactic structure gives rise to only (b).

What is independently needed, regardless of the mode of application, is a notion of the domain of tone sandhi. Given the domain of tone sandhi, the facts become more
comprehensible. In (d) the domain of application is the whole sentence and in the rest of the variants smaller domains are involved.

3.4.2. The domain of Third Tone Rule

There are a number of different positions one can take with respect to the nature of the domain of tone sandhi. The domains can be linear or hierarchical; the domains can also differ depending on whether they are morphosyntactic units or prosodic units.

The following chart gives some representative analyses, grouped along the dimensions of syntactic versus prosodic domains and linear versus hierarchical domains:

<table>
<thead>
<tr>
<th>LINEAR</th>
<th>HIERARCHICAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNTACTIC</td>
<td>Cheng (1973)</td>
</tr>
<tr>
<td>PROSODIC</td>
<td>----</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.4.2.1. Linear versus hierarchical.

In this section, Cheng's analysis, which assumes linear domains, will be examined and argued to be inadequate. A defence of cyclic application on hierarchical domains will be given.

3.4.2.1.1. Cheng (1973) uses flattened syntactic structures as the tone sandhi domains, whose boundaries
correspond to syntactic constituents. The domains are not hierarchical because the rule does not reapply in general.

Cheng uses 'syntactic depth' as the variable by which the size of the domain is determined. 'Syntactic depth' refers to the junctural strengths between constituents. The depth is 1 within the NP, 2 within the VP and 3 for the whole sentence. The greater the depth, the larger the domain. The choice as to the depth at which the rule applies is based on rate of speech. At the slowest rate, the rule applies only in NPs; at a faster rate, the rule applies in the VP domain; at an even faster rate, the rule applies in the domain of the whole sentence.

A problem with Cheng's analysis and theories with linear domains in general is that linear domains cannot accommodate the asymmetry between left- and right-branching structures, which can be captured only in a cyclic approach assuming hierarchical domains. The asymmetry is seen in the following examples:

```
mai hao jiu
buy ASP. wine
'has bought wine'
3 3 3
*2 3 3
2 2 3

mai hao jiu
buy good wine
'buy good wine'
3 3 3
3 2 3
NA
```

* UR

* cycle 1

* cycle 2
At the same rate of speech, the left- and right-branching structures give rise to different outputs, which for convenience will be called alternating and non-alternating patterns hereafter.

3.4.2.1.2. The necessity of cyclic application. In linear domains multiple application has to be used if more than two third toned-syllables are involved. Either simultaneous or left-to-right iterative application can be used, whereas right-to-left is impossible due to the dissimilatory nature of the rule:

**SIMULTANEOUS**

\[
\begin{array}{ccc}
3 & 3 & 3 \\
\rightarrow & 2 & 2 & 3
\end{array}
\]

**LEFT-TO-RIGHT ITERATION**

\[
\begin{array}{ccc}
3 & 3 & 3 \\
\rightarrow & 2 & 3 & 3 \\
\rightarrow & 2 & 2 & 3
\end{array}
\]

**RIGHT-TO-LEFT ITERATION**

\[
\begin{array}{ccc}
3 & 3 & 3 \\
\rightarrow & 3 & 2 & 3 \\
\text{NA} \\
3 & 2 & 3
\end{array}
\]

Simultaneous/iterative application does not respect the internal structures of the domain in which a rule applies. Both left- and right-branching structures give rise to the same result. But cyclic application does respect hierarchical structures. On left-branching hierarchical structures, it generates non-alternating strings. On right-branching structures, though, cyclic
application gives alternating 2s and 3s for a string of underlying third tones.

On the other hand, simultaneous application can generate non-alternating strings of [2 2 3] on right-branching structures in faster speech; and cyclic application cannot. This, in fact, is the reason Cheng gave in rejecting cyclic application.

So far, it seems that cyclic application is as inadequate as simultaneous application in accounting for the whole range of facts, if not more so, due to the extra structure that has to be assumed. But we shall see that whereas hierarchical syntactic structure is independently existent for cyclic application, the mechanism which non-cyclic theories has to use to accommodate the different outputs is rather ad hoc. Furthermore, although simultaneous application can generate both non-alternating strings for both left- and right-branching structures, the uniformity of the mode of application employed is not necessarily desirable, since the non-alternating strings for right-branching structures are possible only in fast speech. On the other hand, although cyclic application cannot generate such strings, this is not necessarily bad, as other modes of application can well be used for fast speech.

In Cheng's approach, the variability in the domain sizes depends on the rate of speech. For the pattern [2 2
3] to be generated on a left-branching structure with simultaneous application, the domain will have to be the whole structure. But in order for [3 2 3] to be generated on a right-branching structure, the domain cannot be the whole string. In other words, the domains for the left-branching structure and the right-branching structure have to be different in size in order for the output to be different. But the left-branching structure, having a larger domain, does not have to be pronounced faster than the right-branching structure.

The argument for cyclic application can then be summarized as follows:

a) Cyclicity can account for the asymmetry between left- and right-branching structures, i.e. the patterns [2 2 3] and [3 2 3] respectively.

Theories not including cyclic rule application cannot account for such asymmetry without positing domains of different sizes for the two kinds of structures, larger for left-branching structures and smaller for right-branching structures. But speech rate, which is appealed to as being responsible for the variation in domain sizes, does not necessarily differ for right- and left-branching structures.

b) Theories not including cyclic rule application can account for the two patterns for right-branching structures [3 2 3] and [2 2 3] by positing different
domains sizes attributable to arguably real difference in speech rate.

Cyclicity can also account for such variation if they are allowed to vary the initial cycle of application depending on the speech rate. Although freely choosing the initial cycle of application seems quite ad hoc, it is no more so than varying the size of the domain, as is done in non-cyclical theories.

We can conclude that cyclic application is superior to simultaneous application. By implication, hierarchical domains are necessary.

3.4.2.2. Syntactic versus prosodic.

Now that cyclic application and hierarchical domains are argued to be necessary, what remains is the choice between approaches positing syntactic domains and those with prosodic domains. The analyses to be reviewed here include Kaisse (1985), which is a syntax-based approach and Liu (1980) and Shih (1986), which are both prosodically-based approaches.

3.4.2.2.1. Kaisse (1985) uses syntactic structures as the sandhi domains, on which the Third Tone Rule cyclically applies. The application of the rule is subject to the Branch Condition, which will not be discussed until later. Kaisse's analysis is based on Liu (1980).
3.4.2.2.2. Liu (1980), adopting the framework of Liberman and Prince (1977) and Selkirk (1978), uses metrical structures as the domain for tone sandhi. Her mode of application is also cyclic. However Liu's metrical structures are not that different from the syntactic structure in configuration, as can be seen from the fact that Liu's analysis is readily translated into Kaisse's analysis. In any case, the algorithm mapping syntactic structures onto prosodic ones is not given in Liu's paper.

Liu's analysis is different from Kaisse's in one respect, though. The metrical trees of Liu have syllables as terminal units rather than words (as in Kaisse's analysis). In this way, polysyllabic morphemes having no morphosyntactic structures can nonetheless have prosodic, internal structures.

3.4.2.2.3. Shih (1986) is a more radical, prosodically based analysis. She explicitly posits prosodic structure that is non-isomorphic with syntactic structures.

Shih differs from previous investigators in discovering the following kind of data and the drastic prosodic approach that is needed to account for them:
In both sentences, [gou yao], a domain of tone sandhi, does not correspond to any syntactic or semantic constituent. Such examples call for a radical prosodic theory.

The mechanism Shih adopts, with some modification, for building prosodic structures as sandhi domains, is Chen's (1979) Foot Formation Rule, originally devised to account for versification. It is given below:

Foot Formation Rule (FFR):

I. Foot (f) Construction
   a. IC (Immediate Constituency): Link immediate constituents into disyllabic feet.
   b. DM (Duple Meter): Scanning from left-to-right, string together unpaired syllables into binary feet, unless they branch to the opposite direction.

II. Super-foot (f') Construction
   Join any leftover monosyllable to a neighboring binary foot according to the direction of syntactic branching.
The sentence 'gou yao xiao hua mao' can be given the following prosodic structure:

```
  s
 /    \\
|     | vp
| np  /    \\
| np  | np
   /    \\
gou yao xiao hua mao
```

Tone sandhi proceeds in the following fashion:

```
P
 /    \\
|    | f
| f'  /    \\
|    | f
|    /    \\
gou yao xiao hua mao
dog bite small flower cat
3 3 3 1 1
2 3 cycle 1
2 2 cycle 2
```

The application within the foot or the superfoot is obligatory; but across feet, application is optional.

These points are illustrated in the following example:

```
lao li yi jing mai hao jiu
old li already buy ASP wine
3 3 3 1 3 3 3 3 UR
[ ]f [ ]f [ ]f by IC
[ ]f' by Super-f
? 2 3 3 1 2 3 3 cycle 1, f
2 3 3 1 2 2 3 cycle 2, f'
2 2 3 1 2 2 3 cycle 3, p
```
3.4.2.2.4. The necessity of independent prosodic structures. It is most important to observe that prosodic structures non-isomorphic with syntactic structures are necessary when we find tone sandhi applying in a domain that corresponds to no morphosyntactic constituent, as the class of examples given by Shih shows.

Independent prosodic structures can also be found in polysyllabic simple words such as loan words and transliterations, in which the rule can apply cyclically. It will be seen later on that the default prosodic structure in trisyllabic simple words is left-branching, giving rise to the expected non-alternating string:

```
[So ma li]m
/ \ /
3 3 3
/ \ /
2 2 3
```

The most serious problem we have seen for the syntactic approach, assuming cyclical application, is that it cannot get non-alternating patterns like [2 2 3] on right-branching structures in fast speech. Shih's solution is to vary the initial cycle of application, i.e., larger for [2 2 3] and smaller for [3 2 3], and to allow simultaneous application for the larger domain. The availability of non-isomorphic prosodic structures opens
up new possibilities for the solution of the problem. Given the possibility of non-isomorphism, the non-alternating strings can now be generated if right-branching structures can be changed into left-branching ones in fast speech.

3.4.3. The Present Analysis.

The present analysis is summarized as follows:

a. The domain of the Third Tone Rule is prosodic;
b. The prosodic domain is hierarchical;
c. The mode of application is cyclic;
d. The construction of the domain is sensitive to the pragmatic, discourse structure;
e. The construction of the domain occurs both lexically and post-lexically;
f. The construction of the domain is subject to the reinterpreted Branch Condition;
g. Reanalysis occurs in special contexts.

As can be seen above, there are a number of features the present analysis shares with previous analyses: in adopting cyclic application and hierarchical domains, I concur with Liu, Kaisse and Shih; the adoption of prosodic domains non-isomorphic with syntactic structures is
similar to Shih. My proposal is similar to Liu's and Kaisse's in including the reinterpreted Branch Condition.

The differences between my analysis and previous ones are the following:

1) the relevance of pragmatics and discourse in constructing prosodic domains;

2) the separation of domain construction into a lexical and a post-lexical part;

3) the relevance of the Branch Condition in domain construction;

4) the necessity of reanalysis.

In the following, I will motivate these differences in turn.

3.4.3.1. The Relevance of Pragmatics and Discourse Contexts. One difficulty in accounting for the Third Tone Rule is the variability in the sandhi output. The variability has mostly been attributed to rate of speech. Cheng's example is given below, with the purported different rates of speech indicated:

<table>
<thead>
<tr>
<th></th>
<th>lao</th>
<th>li</th>
<th>mai</th>
<th>hao</th>
<th>jiu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Li</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>buy</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>good</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>wine</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

> a. 2 3 3 2 3 adagio
b. 2 2 3 2 3 moderato
c. 2 3 2 2 3 allegro
d. 2 2 2 2 3 presto

To be sure, Cheng (1973) mentioned, in addition to rates
of speech, the effect of what he called 'attitudinal factors' in Mandarin Third Tone Rule. He seems to suggest that casualness is one such factor but did not mention what other attitudinal factors are relevant. Nor did he give any account of how attitudinal factors play a role. Later analyses seem to have ignored such factors entirely, with the exception of L. Cheng (1987), who suggested that semantic/pragmatic differences in different renditions of the same sentence should be investigated.

Elicitation of data has also proceeded without consideration of the pragmatic/discourse contexts. The most common procedure is represented by the question posed by the investigator to the informant: 'Can you say this?' The answer is often inconsistent from person to person, from occasion to occasion; sometimes everything seems acceptable. A more sensible approach would be to elicit responses with concrete pragmatic/discourse situations. Instead of 'Can you say this?', the question should be posed as: 'What do you say in this environment?'

In theories without reanalysis, rate of speech can be used as a variable only to control the size of existing domains, as Cheng and Shih did in having larger domains for faster speech. It cannot, however, alter the groupings of syllables. If evidence is found showing domains of tone sandhi distinct from each other as well as from the syntactic structure, something other than rate of
speech would have to be responsible for the difference. As a matter of fact, there is such evidence, as shown in the example that follows:

\[\text{wo xiang mai xiao hua mao} \]
\[\text{I wish buy small flower cat} \]
\[3 \quad 3 \quad 3 \quad 3 \quad 1 \quad 1 \quad \text{UR} \]

a. 3 \ [2 \ 3] \ 3 \ 1 \ 1 \ \text{revised FFR?} \\
a'. 3 \ [2 \ 2] \ 3 \ 1 \ 1 \ \text{revised FFR?} \\
b. 2 \ 3 \ 2 \ 3 \ 1 \ 1 \ \text{cycle on syntax} \\
c. [2 \ 2] \ 3 \ 3 \ 1 \ 1 \ \text{FFR(Shih)} \\
c'. [2 \ 2] \ 2 \ 3 \ 1 \ 1 \ \text{FFR(Shih)} \\

There are at least three groups of variants, the (a) and (a'), (c) and (c') being further variants within a group due to the optional application in (a') and (c'). Of these groups, only the (b) variant can be generated by cyclic application on syntactic structure. Since the syntactic structure is right-branching all the way, the tonal pattern is alternating. While (b) is a possible reading of the sentence, it is by no means the unmarked reading. To say the least, the other two groups are also possible. But since they cannot be generated by cyclic application on syntactic structure or by taking the initial cycle to be on larger domains isomorphic with syntactic structures, their tonal domains have to be constructed more independently. Shih's Foot Formation Rule can give only the (c) group:
The (b) reading is not allowed by Shih, because the third and the fourth syllables cannot form a foot by DM, due to the 'lexical integrity principle' protecting xiao from being reanalyzed out of 小花毛, a lexical item for her. The (a) group seems to be the unmarked reading; but it cannot be generated by either cyclic application on syntactic structure or by Shih’s FFR.

The different forms are not free variants, though. The fact that native speakers consent to all these forms seems to create the impression that they are. But in almost all elicitation sections, the question posed to the informant is 'Can you say this' rather than 'Can you say this in this context' or 'What do you say in this context'. In the absence of the control for pragmatic contexts, native speaker intuition also tends to be inconsistent. Once we take pragmatics into consideration, the judgments become sharper.

Let us examine the same sentence again. The (a) group, the most unmarked reading, can be said with minimal context or as part of the following exchange:
A: Ni mai xiao hua mao le ma?  
'Have you bought the small flower cat?'

a. 3 2 3 3 1 1 
a'. 3 2 2 3 1 1 
B: Mei ne. Wo xiang mai xiao hua mao. 
'Not yet. I am thinking of buying the small flower cat'.

When the sentence occurs in the above exchange, there is contrastive stress on xiang. But in the most unmarked reading with minimal context, there is no contrastive stress on the word. The reason why the two readings are tonally identical can be explained by saying that in the most neutral reading, the new information is conveyed in the verb phrase; and since xiang is the head of the verb phrase it also receives the sentential stress.

The (b) reading sounds most natural in the following exchange:

A: Ni xiang mai xiao hua mao ma?  
'Do you wish to sell the small flower cat?'

b. 2 3 2 3 1 1 
B: Bu. Wo xiang mai xiao hua mao.  
'No. I wish to buy the small flower cat.'

The contrast is between mai 'sell' with the 4th tone and mai 'buy' with an underlying 3rd tone, which happens to be segmentally identical. There is also contrastive stress on the word mai 'buy'.

The (c) group occurs most naturally in the following
exchange:

A:  Ta xiang mai xiao hua mao ma?
    'Does he wish to buy the small flower cat?'

c.  2  2  3  3  1  1

B:  Bu.  Wo xiang mai xiao hua mao.
    'No.  I wish to buy the small flower cat.'

The contrastive stress is on the first syllable wo 'I'.

Clearly, much of the variation in the Third Tone Rule is pragmatically conditioned. How do we construct the sandhi domain then? A uniform algorithm like that of Shih's clearly will not do. Following are examples showing that many of the variants, often unmarked, are not predicted by her analysis.

wo  lao xiang mai shui guo
I always wish buy water fruit
'I always want to buy fruit'.
3  3  3  3  3  3

by IC

[a:  2  3  2  3  2  3  predicted

b.  3  2  3  3  2  3  unpredicted

b'.  3  2  2  3  2  3  unpredicted

The predicted reading is not the unmarked one because it sounds natural only when either wo 'I', or xiang 'wish' has contrastive stress. The unpredicted readings are either unmarked or has emphasis on lao 'always'. Another example is given below:
wo mai shui gang
'I buy water jug'
3 3 3 1
[  ]f by IC
[  ]f by DM
a. 2 3 3 1 predicted
a'. 2 2 3 1 predicted
b. 3 2 3 1 unpredicted

Again, the unpredicted reading is the unmarked one.

The examples above show that Shih's DM as stated excludes some unmarked readings. For ease of reference, the DM is repeated below:

DM (Duple Meter): Scanning from left-to-right, string together unpaired syllables into binary feet, unless they branch to the opposite direction.

The inability to accommodate different variants lies in the decision to scan from the first syllable of the utterance in all cases, regardless of semantic/discourse structure.

The hypothesis I put forward is the following:

DM application:

The DM starts scanning from the syllable that bears the sentential accent. In unmarked cases, the sentential accent falls on the head of the verb phrase. Otherwise, the sentential accent falls on the syllable bearing contrastive stress. In case the accented syllable is other than the first syllable, apply the DM again from the beginning of the sentence.
The following sentence illustrates the various points at which DM starts scanning:

\[
\begin{array}{cccccc}
\text{wo} & \text{xiang} & \text{mai} & \text{xiao} & \text{hua} & \text{mao} \\
\text{I} & \text{wish} & \text{buy} & \text{small} & \text{flower} & \text{cat} \\
3 & 3 & 3 & 3 & 1 & 1 \\
\end{array}
\]

a. \(3 2 3 3 1 1\) DM, 2nd syl.
a'. \(3 2 3 3 1 1\) DM, 2nd syl.
b. \(2 3 2 3 1 1\) DM, 4th, 1st syl.
c. \(2 2 3 3 1 1\) DM, 1st syl.
c'. \(2 2 2 3 1 1\) DM, 1st syl.

Let us go through the derivation for each reading. First, the derivation of (c) and (c') is given below:

\[
\begin{array}{cccccc}
\text{wo} & \text{xiang} & \text{mai} & \text{xiao} & \text{hua} & \text{mao} \\
\text{I} & \text{wish} & \text{buy} & \text{small} & \text{flower} & \text{cat} \\
3 & 3 & 3 & 3 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c}
[ ]f \\
\text{by IC} \\
[ ]f' \\
\text{by DM, 1st syl.} \\
[ ]f'' \\
\text{by super-f} \\
c. \ 2 & 2 & 3 & 3 & 1 & 1 \\
\text{within f'} \\
c'. \ 2 & 2 & 2 & 3 & 1 & 1 \\
\text{between f'}
\end{array}
\]

Both pronunciations, derived from applying DM from the first syllable, are appropriate responses to the question "Does he wish to buy a small flower cat?"

Two details of derivation need to be explained here. Shih would not group the fourth syllable xiao with the preceding syllable mai on account of the claim that xiao is part of the lexical item formed by the last three syllables, and hence is not prosodically reanalyzable. Thus the third syllable mai has to be grouped into a
superfoot with the preceding foot. But as it now stands, there will be no way in Shih's account to do this, as mai is in a right-branching structure and can group only with things occurring to the right. But since to the right of mai there is already a superfoot it cannot take mai. To account for (c), a stray syllable will be subsumed under a superfoot with either the preceding or the following foot regardless of the syntactic configuration in which the stray syllable occurs.

The derivation of (b) is given below:

```
wo xiang mai xiao hua mao
I wish buy small flower cat
3 3 3 3 1 1

[ ]f by IC

[ ]f by DM, 4th

[ ]f by DM, 1st

b. 2 3 2 3 1 1
```

This is the reading where mai 'buy' contrasts with mai 'sell'. It is derived by two applications of DM, once from the 4th syllable, once from the 1st syllable. The pronunciation, while no different tonally from the one derived by applying DM from the first syllable alone, is different in accent by having it on mai rather than elsewhere.

Finally, the readings (a) and (a') are given below:
This is the reading where xiang 'wish' is the accented syllable.

Since the accented syllables always start the sandhi domain, they have the changed tone, i.e. the second tone. This might seem counterintuitive, because we normally expect syllables with emphasis or contrastive stress to retain their original tones. But stress and the changed tone do cooccur, as evidenced by the examples above.

Confirmation of this cooccurrence can be found in the cleft construction, which is another way to bring material into focus:

\[
\text{shi wo mai de jiu}
\]
\[
\text{is me buy prt. wine}
\]
\[
'\text{It's me who bought the wine'}
\]

It is unproblematic that the first reading is accepted; it is unclear whether the second one, which retains the original tone on the syllable with emphasis, is acceptable or not.
3.4.3.2. Lexical Foot Formation. We have thus dissected the mechanism of DM into two parts, one applying from the syllable bearing the sentential accent, one from the beginning of the whole sentence. From what we have said so far, there seems to be nothing special about the formation of sandhi domains in lexical items. Foot Formation can just as well be applied to lexical items:

```
  so   ma   li
  'Somali'
  3    3    3    by DM
[   ]f    by super-f
[   ]f' by super-f
2    3    3    1st cycle
2    2    3    2nd cycle
```

Shih thus concludes that there is only one Foot Formation Rule. If lexical items are given in citation form, as Shih does, there is indeed no problem, as we can always start scanning from the first syllable and get the desired result. A problem arises when we embed lexical items in sentences and apply domain construction at the sentence level. There is then no guarantee that scanning will always start from the beginning of the lexical item. We may, for example, find the following situation, when the last two syllables of the lexical item are grouped together and the first syllable is grouped with something else outside the lexical item:
wo hen xiang suo ma li
I very miss Somali
'I miss Somali very much'

According to Shih, such cases are prohibited by the Lexical Integrity Hypothesis (LIH). But the LIH is a negative constraint only. Assuming the LIH, Foot Formation has to blindly generate prosodic structures until it gives the right one. Alternatively, Foot Formation has to look ahead in order to generate the right structure and thus acquires global power.

I thus modularize Foot Formation further:

Apply Lexical Foot Formation first in the lexicon.

I choose to apply the whole Foot Formation Rule in the lexicon rather than just the IC and DM components because the super-foot cannot be violated either by domain formation at the post-lexical level. In the following example, shui guo jiu 'fruit wine' is a lexical item and the superfoot formed cannot be violated at the post-lexical level:
The syllable jiu is grouped into a super-foot with the preceding foot:

```
shui guo jiu hao
water fruit wine good
'Fruit wine is good'
3 3 3 3
*[ ]f [ ]f by DM
*2 3 2 3
```

Applying FFR in the lexicon eliminates the need for the LIH, since the prosodic structure formed in the lexicon will not be altered later on, except in fast speech.

Applying FFR in the lexicon is supported by the lexical application of the Third Tone Rule. While there is no a priori reason that the lexical application of the rule has to occur in the same domains as post-lexical application, such an assumption is the default one. The evidence that the Third Tone Rule applies in the lexicon is given below:

```
3 3 ----> 2 3 ----> 2 N
lao hu
old tiger 'tiger'
```
3 3 ———> 2 3 ———> 2 N
xiao jie
small sister 'miss'

The output of the Third Tone Rule undergoes a lexically
governed rule of tonal deletion (The N means atonic).
Since the tonal deletion rule is lexical, the Third Tone
Rule, which feeds it, has to be lexical as well.

3.4.3.3. Revised Foot Formation Rule. We posit the
following hypothesis concerning the application of Foot
Formation Rule, in particular, DM:

  FFR applies in the lexicon first;
  DM applies from the syllable with sentential accent;
  DM applies from the first syllable of the whole
sentence;
  Super-foot formation applies to the whole sentence.

3.4.3.4. The Relevance of the Branch Condition.

  In her analysis of Mandarin Third Tone Rule, apart
from using syntactic structures as the cyclic domains,
Kaisse also posited the Branch Condition, which is a
constraint on the application of the rule based on
syntactic configurations. It was proposed originally by
Liu (1980) in her metrical treatment but was restated in
syntactic terms by Kaisse:
Tone sandhi may apply between two words a and b if a is the left branch of the constituent that contains b or if b is the right branch of the constituent that contains a; in other words, tone sandhi applies if the sandhi pair is on an edge of the constituent that contains it.

The following shows two of the possible configurations allowing sandhi between a and b:

```
    / \  / \  \
   / \ / \ / \
  a  b a  b
```

In fact, the condition should be stated more strongly than the may in the original formulation suggests; sandhi between a and b must apply in the above configurations, as evidenced by the following two examples:

```
    / \  / \  \
   / \ / \ / \
  wo mai shu lao li hao
 I buy book old Li good
'I buy books' 'Old Li is good'
 3  3  1   3  3  3 UR
 *3  3  1 *2  3  3 cycle 1
 2  3  1   2  2  3 cycle 2
```

Although the condition does not explicitly state it, it is supposed to rule out the application of the rule between a and b in the following configuration:

```
    / \  \
   / \  \
  c  d
    / \  \
   a/b
```
The Branch Condition then is used to capture a real contrast between tone sandhi pairs lying at the edges as opposed to non-edges. It is not clear though that the contrast is between application and non-application. While application at edges is obligatory, at non-edges application is not impossible, but rather optional:

```
lao li mai hao jiu
old li buy good wine
3 3 3 3 3
2 3 3 2 3
 obligatory
2
--------------
2 2 3 2 3
 optional
```

Given this fact, the Branch Condition can perhaps be interpreted as making sandhi obligatory for pairs lying at the edges but optional for those at non-edges.

An important proviso that should be stated is that obligatoriness does not mean sandhi will in fact occur but only that it will occur when the environment for it is not destroyed by rule application at an earlier cycle. For example, in the following right-branching structure, the first two syllables cannot undergo sandhi even if they lie
There is a possible use of the Branch Condition in a theory assuming prosodic domains, which is to make edges possible positions for domain formation and/or non-edges impossible for domain formation. Application of the rule within a domain is obligatory but it is optional across domains.

The Branch Condition then captures a real contrast between edges and non-edges. The contrast can be one between obligatory and optional application or one between possible sandhi domain and impossible ones.

Shih explicitly argued against the Branch Condition of Liu and Kaisse, giving examples of sandhi occurring where the condition forbids it. If taken as argument against the letter of the condition, the criticism is basically correct; but as the spirit of the condition is to capture the contrast between edge and non-edge positions, whether it is a contrast in the optionality of rule application or one in domain formation, her argument no longer is well-directed. We will see that some form of the condition is necessary even for Shih's analysis.
We will look at the most important part of Shih's proposal, the Duple Meter part of the Foot Formation Rule.

Scanning from left-to-right, string together unpaired syllables into binary feet, unless they branch to the opposite direction.

A careful reading reveals that some aspect of the Branch Condition is inherent in it. The thing to notice here is the conditional clause at the end. Shih specifically brought attention to the fact that this conditional clause is not present in Chen's original formulation in the context of poetic scansion (1979). She added this clause to account for the difference between natural language and poetry. She was right in noticing the difference between poetry and natural language. She failed to notice that the conditional clause is basically the Branch Condition reinterpreted! Note the similarity in prediction between the Branch Condition and the exceptional clause of DM:

Branch Condition: \[ \frac{a/b}{/=no \ (obligr.) \ sandhi} \]

Condition on DM: \[ *[a\ b]f \ *[\ ]=no \ foot \ formed \]

In neither case do we get obligatory sandhi between a and b. This is predicted by the Branch Condition, if we reinterpret the contrast between edge and non-edge positions as one between obligatory and optional
application rather than that between the occurrence and non-occurrence of sandhi; it is predicted by the restriction on DM too, because obligatory foot-internal sandhi is impossible but optional, inter-feet sandhi is still possible.

From the discussion above, we can see that although syntactic configuration is not directly regulating tone sandhi, it nonetheless plays a role in the construction of prosodic domains. Therefore, we acknowledge the role of the Branch Condition thus:

Revised DM:

Duple Meter: Scanning from left to right, group together any stray syllables into disyllabic feet, subject to the restriction imposed by the Branch Condition.

The Revised Branch Condition, while similar in spirit to Kaisse's version, is restated in terms of the possible formation of prosodic foot:

Revised Branch Condition:

Two syllables cannot form a prosodic foot if neither of them lies at the edge of the constituents that contain them.

For example, in the following tree, a and b do not form a prosodic foot:

```
/ \  
/ \  / \  
*[a b]f
```
3.4.3.5. The Necessity of Reanalysis. The application of the Third Tone Rule can be further complicated by the overriding of the IC by DM. Such overriding occurs only in special circumstances such as fast speech. Therefore we cannot simply eliminate regular prosodic domain formation, which includes IC as a component. A more constrained way to handle the situation will be to posit reanalysis in special circumstances. Reanalysis in this particular instance is not ad hoc, as it is independently needed for the application of another tonal rule in the language. Reanalysis also offers an analysis of the super-foot that is more coherent with the rest of the theory in terms of the mode of application.

3.4.3.5.1. Conflict between DM and IC. According to FFR, the application of DM follows that of IC and applies to syllables left over from IC. Therefore DM should not be able to alter the structures constructed on the basis of IC. There are examples, however, showing the overriding of IC by DM:
In the first two examples, the one form generated by using the IC foot as domain is questionable. In the third example, we have to explain the acceptability of the form generated by using the DM foot as domains. Since the domains constructed on the basis of IC and DM are mutually exclusive, we will have problem accounting for the existence of either of the forms in any case.

In the face of such examples, we have to concede that prosodic structures are not constructed with one mechanism alone but by two separate ones. The two mechanisms are:

1) initial prosodic structure construction like FFR, used in normal conditions;
2) restructuring under special circumstances.

3.4.3.5.2. Analysis of Super-foot. The super-foot is a tri-syllabic or quadrisyllabic unit of various configurations. In a right-branching superfoot, an analytical problem presents itself. In normal speech, sandhi occurs only between the second and the third syllable, i.e. in the foot constructed based upon IC constituency. This is handled well with cyclic application. But in faster speech, sandhi includes the first syllable as well.

mai hao jiu
buy good wine

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>UR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[</td>
<td>]f</td>
<td>by IC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[</td>
<td>]f'</td>
<td>by super-f</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>3</td>
<td>normal speed</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>fast speech</td>
<td></td>
</tr>
</tbody>
</table>

This can be handled by only simultaneous application or left-to-right iteration, given the right-branching structure. What Shih did is to postulate the foot as the domain of sandhi at normal speed but at faster speech the superfoot as the domain where simultaneous application is used.

I suggest that a more coherent theory will result if cyclic application is adhered to in cases where Shih
adopts simultaneous application. All the problematic cases involve non-alternating tonal patterns on right-branching structures. Left-branching structures, on the other hand, are perfectly amenable to cyclic application. It is only natural then that Poteet (1985) came up with a proposal for 'reanalysis', which changes right-branching structures to left-branching ones.

Shih considers Poteet's reanalysis ad hoc. In the absence of independent evidence, such a solution indeed appears less than motivated. I will point to independent evidence, however, supporting such an analysis. The evidence comes from the prosodic domain for the application of the Second-to-First Tone Rule.

It was pointed out in the discussion on the Second-to-First Tone Rule that in trisyllabic forms with final full-toned syllables, the first two syllables are grouped into a prosodic constituent in fast speech, as is shown below:

\[
\begin{array}{c}
\text{[ [ S S ] S ]}
\end{array}
\]

The existence of this prosodic constituency is supported by the lenition and resyllabification facts, examples of which are repeated here:

\[\begin{array}{c}
a. \quad \text{bai} \quad \text{shi} \quad \text{qiao} \\
-----\rightarrow \quad \text{bai .} \quad \text{ri .} \quad \text{qiao} \\
\quad \text{white stone} \quad \text{bridge}
\end{array}\]
We thus have independent evidence of reanalysis in fast speech. Given the left-branching structures that result from reanalysis, all we have to do, in order to get non-alternating tonal patterns, is to let the rule apply cyclically on prosodic structures.

This account is more elegant than Shih's super-foot account, since it makes use of the already needed cyclic application rather than adding simultaneous application specifically for the level of super-foot. Although reanalysis seems a complication, it is independently needed in the application of the Second-to-First tonal rule.

3.4.3.6. The role of prosody, morphosyntax and discourse

From the Foot Formation Rule formulated by Shih, we see that both morphosyntactic structures and rhythmic patterns of the language contribute to the prosodic structure that ultimately result. This is seen in the fact that both IC and DM are necessary in the algorithm. The FFR also implicitly predicts that IC takes precedence
over DM, as IC is ordered before DM.

In the present analysis, the roles of morphosyntax and rhythmic patterns are more explicitly differentiated by incorporating the revised Branch Condition into the DM and the separation of FFR into a lexical and a post-lexical component.

The present proposal also incorporates the role of pragmatics and discourse contexts in the construction of prosodic domains, resulting in a clearer account of the variability in the output of the Third Tone Rule.

Reanalysis alters prosodic structures in favor of the underlying rhythmic patterns of the language in special circumstances. One such context is poetry, which places a high premium on strict rhythmic patterns. Here, the Branch Condition no longer applies and DM can join two syllables that branch in different directions in the syntactic configuration (Chen 1979). In poetry, too, the IC part of the FFR may no longer take precedence over DM.

Another context for reanalysis is fast speech. Here the effect of IC diminishes and that of DM increases, resulting in reanalysis converting right-branching structures to left-branching ones.
Notes to Chapter 3

1. This needs to be investigated more carefully. Although the variant pronunciation of san may be due to dialectal mixture, there is always the possibility that it has not been noticed before as a variant form in Mandarin proper. Lin (1985), which brings attention to the hitherto unreported change of a 4th tone to a 2nd tone before another 4th tone, shows that this is a distinct possibility.

2. Such a rule, if it exists, would be very restricted in its application.

3. Although Yip made reference to word trees, it seems that she meant phrase trees, as she denied that there was any distinction in Mandarin between word and phrase in terms of stress. From Yip's later work (1982) on Mandarin stress, it seems that she later entertained a separate word level.

4. Margie Chan pointed out that M. Chen (1979) has a different algorithm for assigning stress values, according to which the sentence has increasingly strong stress.

5. It is not clear to me at this point what this level should be. The possibilities are the word or the foot.

6. The tree is drawn here assuming that it is analogous to the consistently right-branching tree for Lao Li mai hao jiu 'Old Li buys good wine', which Yip gives in her dissertation.

7. See chapter 6 for a discussion of the possible differences between the prosodic category serving as the domain for the Second-to-First Tone Rule and the one for the Third Tone Rule.
8. Most recent analyses of Mandarin assume that the third tone is underlyingly low-level. I showed in 3.2 in connection with the Half Third Tone Rule that the assumption should be examined anew.

9. Cheng allows reapplication only for sandhi pairs at the edges of the sentence. In any case, reapplication does not follow morphosyntactic bracketing.

10. In Chinese tone sandhi literature, the correlation of the sandhi position with the weak position has been frequently noted.
CHAPTER IV
DOMAIN OF TONE IN DANYANG

In this section, Danyang, a northern Wu language, will be examined. Danyang is spoken in Danyang county in northern Jiangsu province. Its tone sandhi facts are described in detail in Lyu (1980). Danyang has recently attracted attention in connection with the issue of contour tones. Chen (1986) argues that in Danyang a particular contour tone 42 spreads as a unit, giving rise to 42 42 42 in multisyllabic forms. Yip (1988) also makes use of the fact in arguing for the unitary nature of contour tones at some level of the grammar. Danyang is interesting here because of the domain of melodies and its prosodic organization in general.

It will be shown that in Danyang, the domain of tones is phonological rather than morphosyntactic. Furthermore, the phonological unit that bears melodies in Danyang is the foot. The construction of the foot is rather similar to that in Mandarin, i.e. syllables are grouped into disyllabic feet starting at the beginning of a domain. In fact, the tendency to favor left-branching prosodic structure is more clearly manifested in Danyang.
than in Mandarin.

4.1. Melodies in Danyang

Danyang is different from Mandarin in one important way, that is, the melodies on polysyllabic expressions cannot be taken as the function of the component syllable melodies. In fact, the melodies on polysyllabic expressions are neither derivable from synchronic citation tones nor according to historical tonal classes. In order to present the argument, an introduction to the historical tonal categories and their relationship with modern tones will be given first.

4.1.1. Historical tonal categories.

Historically, Middle Chinese has four tone categories (syllable types\(^1\) incorporating both segmental and tonal contrasts) called ping, shang, qu, ru, the first three non-checked and the last one checked (ending with stops). Each of these categories later split into two subcategories, conditioned by the initial consonant/phonation types. In other words, a register distinction arose. Those with initial voicing/murmur are called 'yang', those without 'yin'. Those with final stops are called 'entering' tones. The eight types are listed below with their traditional names.
'h' denotes murmur; X stands for nasal segments and glides and K stands for stops \(p, t, k\). In the modern languages, the number of syllable types is generally reduced (with some exceptions). Different dialects, however, have different developments, some of them quite complicated. Following is a diagram showing the relationship between historical categories and tones in a modern language, Shanghai. (For visual clarity, A, B, C and D are used for historical tonal categories, instead of the Roman numerals used above. The a and b indicate upper and lower register respectively).

Middle Chinese:  

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
</table>

Intermediate stage:  

<table>
<thead>
<tr>
<th>Aa</th>
<th>Ab</th>
<th>Ba</th>
<th>Bb</th>
<th>Ca</th>
<th>Cb</th>
<th>Da</th>
<th>Db</th>
</tr>
</thead>
</table>

Modern Shanghai:  

<table>
<thead>
<tr>
<th>A</th>
<th>Cb</th>
<th>Ca</th>
<th>Cb</th>
<th>Ca</th>
<th>Cb</th>
<th>Da</th>
<th>Db</th>
</tr>
</thead>
</table>

We can see from the above that in Shanghai the development takes the form of merging: Ab and Bb have merged with Cb; Ba merged with Ca. No category in the intermediate stage was split into different modern categories.

The situation in Mandarin is different. Not only do we find loss of some categories; for example Da, Db, and
merge of categories, Ca, Cb; we also find that the same category at the intermediate stage was split and distributed among different modern categories. Da corresponds to A, B and C.

Middle Chinese: A B C D
/ \ / \ / \ / \ Intermediate stage: Aa Ab Ba Bb Ca Cb Da Db
| | | | | | | | Modern Mandarin: Aa Ab B B/C C C A/B/C C/Ab

4.1.2. 'Sensitivity to Historical Tonal Categories'.
There is now a consensus among Chinese dialectologists, who generally have extensive field experience with modern dialects and good knowledge of historical phonology, that tone sandhi often is 'sensitive to historical tonal categories'. By historical categories, they either mean the four categories of Middle Chinese, i.e., ping 'even', shang 'ascending', qu 'departing', ru 'entering', or the eight categories resulting from the register split. This realization serves both to guide efforts away from attempts to seek synchronic, phonetic motivation for all tonal alternations and also offers hope of explanation of phenomena that are otherwise utterly mysterious. The sensitivity to historical tonal categories is reflected in two ways: first, syllables with the same phonological tones behave differently in tone sandhi if they belong to
different historical tonal categories; second, syllables with different phonological tones nonetheless behave in parallel fashion if they belong to the same historical tonal categories.

The historical tonal classes need not be restricted to the four historical tonal categories. On the one hand, there can be larger categories formed out of the four basic categories. For example, the categories without stop endings form the category of ru-sheng while the rest constitute shu-sheng; ping-sheng itself forms a class while the other three categories form the class of ze-sheng. On the other hand, within a tonal category, different classes can result based on the types of initials and phonation, the most common division being the one between upper and lower registers.

The sensitivity to historical categories in the tone sandhi of some Wu languages was clearly demonstrated by Ballard (1980). In Danyang as described by Lyu (1980), too, the effect of historical tonal classes is clearly shown.

4.1.3. Danyang Tones.
4.1.3.1. Monosyllabic forms. In colloquial Danyang, six tonal melodies are observed in isolation monosyllables. The six melodies are given below, together with example words (the transcription in the cases of non-checked tones
is according to the Pinyin romanization of Mandarin pronunciation, as Lyu does not give transcriptions to most examples. In the cases of the checked tones with glottal stop ending, Mandarin will not do here. Fortunately, transcriptions for a few checked-tone words can be found:

a. 33 nan 'south'
b. 24 cha 'tea'
c. 55 hao 'good'
d. 11 man 'slow'
e. 3 i? 'one'
f. 4 ia? 'medicine'

Double digits are used for four of the six types even for level tones to indicate longer forms than those with one digit, which are shorter syllables with the glottal stop ending. Ignoring the length difference in the syllables, we find that there are only five contours, which can be given as follows (double digits are used only in the case of contour tones):

\[
\begin{array}{cccccc}
3 & 24 & 5 & 1 & 4 \\
\hat{s} & \hat{s} & \hat{s} & \hat{s} & \hat{s} \\
C V(X) & C V(X) & C V(X) & C V ? & C V ?
\end{array}
\]

Since the last contour denoted by '4' occurs on checked syllables only, we can take it as a variant of the contour denoted by '24', as Lyu did. Hence, there will only be four distinct contours in the monosyllables.
4.1.3.2. Combination Forms. There are six melodies in sandhi forms. They are given below (again, the length differences in syllable types are ignored. So the digit '3' stands for the contours on both a non-checked syllable and a checked syllable; the digit '24' stands for '4' as well on a checked syllable):

There are two things worth mentioning here. When we compare the melodies for disyllabic forms and those of longer ones, we notice a striking correspondence in the shape of the contours. Except for type 3, all the others can be accounted for by spreading of the tones on the
second syllable.

The tone on the second syllable of disyllabic forms also deserves attention\(^3\). In (1), (4) and (6), the second syllable is identical in tone to the preceding syllable; in (2) and (5), the tone on the second syllable seems to be an extension of the tone on the first syllable too; (3) again is an exception. It therefore seems that in all cases but (3), it is the first syllable that provides the contour for polysyllabic forms. If this is the case, there would not be any need to take the disyllable as the basis for melodies on longer forms. Melodies on all polysyllabic forms, including disyllabic ones, would be derived from melodies on the first syllable. While this is a distinct possibility, in the absence of instrumental evidence which bears it out, disyllabic melodies will still be taken as the basis for longer melodies in the following discussion.

4.1.4. Relationship between monosyllables and disyllables.

The relationship between the tones on monosyllabic citation forms and those on disyllables is not as straightforward as that between disyllabic forms and longer forms. Except in verbal reduplication, where the reduplicated forms are mostly derived by spreading the citation tone of the first syllable rightward, the tonal contours of the polysyllabic forms cannot be derived from
citation tones in a straightforward fashion. For example, monosyllables with the same citation tones do not give rise to identical contours on polysyllabic forms:

\[
\begin{align*}
\text{hong} \quad 'red' & \quad \overset{24}{\downarrow} \quad \overset{24}{\downarrow} \\
\text{lao} \quad 'old' & \quad \overset{\n}{\downarrow} \quad \overset{\n}{\downarrow} \\
\text{hongbu} \quad 'red cloth' & \quad \rightarrow 24 \; 5 \\
\text{laobu} \quad 'old cloth' & \quad \rightarrow 3 \; 3
\end{align*}
\]

On the other hand, monosyllables with different citation contours may give rise to the same contour on polysyllabic forms:

\[
\begin{align*}
\text{hong} \quad 'red' & \quad \overset{24}{\downarrow} \\
\text{nan} \quad 'south' & \quad \overset{3}{\downarrow} \\
\text{hongmen} \quad 'red door' & \quad \rightarrow 5 \; 5 \\
\text{nanmen} \quad 'south door' & \quad \rightarrow 5 \; 5
\end{align*}
\]

4.1.5. The insufficiency of historical explanation.

Lyu attributes the above phenomenon to the effect of historical tonal categories. According to him, although hong and lao have the same citation tonal contour, they behave differently in sandhi situations because they belong to different historical tonal categories; and hong and nan behave the same way in sandhi situation despite their difference in citation tone, because they belong to the same historical tonal class. Although historical
explanation works in this case, it may not work in all cases.

Let us now look at the historical tonal classes that Lyu refers to in Danyang. The correspondence between historical tonal categories and modern Danyang citation contours is as follows:

1. Ia. (yin ping) 33
2. Ib. (yang ping) 24
3. IIa. (yin shang) 55
4. IIb. (yang shang) 24
5. IIIa. (yin qu) 24
6. IIIb. (yang qu) 11
7. IVa. (yin ru) 3
8. IVb. (yang ru) 4

As can be seen, quite a few distinctions have merged in the modern language. Historical tonal classes posited by Lyu go beyond the eight classes here though. Instead of the only (a) and (b) register distinction in the four tonal categories, there are four distinctions for each tonal category, depending on the type of initial of the historical syllables. The sixteen classes resulting from this further division are given below, together with the modern isolation tones:

<table>
<thead>
<tr>
<th></th>
<th>clear</th>
<th>half-muddy</th>
<th>yu</th>
<th>full-muddy</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. (ping)</td>
<td>3</td>
<td>3</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>II. (shang)</td>
<td>5</td>
<td>5</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>III. (qu)</td>
<td>24</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IV. (ru)</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

'Clear', 'muddy' 'half-muddy' and 'yu' are terms that are
used in traditional Chinese phonology to differentiate the initial types/phonation types. The phonetic interpretation of these terms is not easy, as there is little indication in the traditional literature as to the phonetic quality of the distinction. 'Clear' refers to 'voiceless obstruents'; 'muddy' refers to 'voiced' or murmurs phonation; 'half-muddy' refers to 'sonorants' and 'yu' refers to 'zero initial'. For the purposes of the present discussion, however, an exact interpretation is not important.

Using the historical affiliation of modern Danyang syllables in terms of these sixteen classes, Lyu tries to predict the tonal contours on sandhi forms. But as we shall see, the prediction is less than perfect. Even with the same historical tonal classes, we still get alternative sandhi patterns. As a dramatic example, there are five tonal patterns resulting from combining syllables that have 3 3 as the first citation tone, of the class ping clear, and 24 as the second citation tone, in the class of qu clear:

\[
\begin{align*}
  &3 &24 \\
\rightarrow &5 &5 \\
&jing &xi \\
&'Peking opera' \\
\end{align*}
\]

\[
\begin{align*}
  &3 &24 \\
\rightarrow &42 &1 \\
&jin &kuang \\
&'gold mine' \\
\end{align*}
\]
The example is extraordinary because of the six tonal melodies that are available for disyllabic forms five are exemplified. That means almost all tonal patterns are realized on syllables with the same citation tones and the same historical class⁴.

The above examples all involve words with different morphemes with different segmental make-ups. A possibility that cannot be ruled out in accounting for the alternative forms is that some as yet undiscovered fact about the segmental composition of the morphemes is at work. Such a possibility can be ruled out on the basis of the following examples:

\[
\begin{align*}
3 & \quad 3 \\
\rightarrow & \quad 3 & \quad 3 \\
\text{kong qi} & \quad \text{'air'}
\end{align*}
\]

\[
\begin{align*}
3 & \quad 24 \\
\rightarrow & \quad 42 & \quad 24 \\
\text{shuang jiang} & \quad \text{'frosting'}
\end{align*}
\]

\[
\begin{align*}
3 & \quad 24 \\
\rightarrow & \quad 24 & \quad 5 \\
\text{beng dai} & \quad \text{'bandage'}
\end{align*}
\]

\[
\begin{align*}
3 & \quad 3 & \quad 3 \\
\rightarrow & \quad 42 & \quad 1 & \quad 24 & \quad 5 \\
\text{gong gong} & \quad \text{gong gong} & \quad \text{'husband's father' 'mother's father'}
\end{align*}
\]

\[
\begin{align*}
24 & \quad 1 & \quad 24 & \quad 1 \\
\rightarrow & \quad 3 & \quad 3 & \quad 1 & \quad 1 \\
\text{bian dan} & \quad \text{bian dan} & \quad \text{'thousand years' egg'}
\end{align*}
\]
Therefore, the variability in tonal patterns cannot always be accounted for with historical tonal classes of the individual syllables. Melodies on some polysyllabic forms cannot be predicted from the component syllables.

4.1.6. Restriction of melodies. Another argument against taking the syllable as the melody-bearing unit is the restriction on possible melodies in polysyllabic forms. The argument can be given either assuming that synchronic citation tones are being combined or taking the tones in combination to be assigned to syllables based on historical categories.

It is easy to see that melodies on synchronic citation tones do not freely combined into polysyllabic forms. The melody 24, which is found on citation forms, cannot be found on any but initial syllables in polysyllabic forms. The tonal values 1, 3 and 5, which all occur on citation tones, do not give rise to any combinations other than 1 1, 3 3 and 5 5 on disyllabic forms.

If we assume that tones are assigned to component syllables/morphemes based on historical classes, we still
predict many more melodies than the actually occurring ones. From discussion in the last section, we found that even when the historical tonal classes involved are exactly identical, we still find a range of tonal values possible for a syllable of a particular position. We noted that disyllabic forms all having 3 as the first citation tone of the class ping clear, and 24 as the second citation tone in the class of qu clear still give \{5, 42, 3, 24, 1\} as the set of values possible for the initial position and \{5, 1, 3, 24\} as the set of values possible for the second position. If we say that melodies on polysyllabic forms are a function of the component syllable/morpheme melodies, we would have to say that the ranges of tonal values should be able to freely combine in disyllabic forms. But in actual fact, a 1 can only follow a 42 and a 1; a 42 can only be followed by either a 1 or 24, and so on. Such mutual dependency in tonal assignment strongly argues against assigning tones to the components.

The only way to make sense of the facts of Danyang is to say that melodies on polysyllabic expressions cannot be derived either from synchronic tones of individual syllables or predicted with reference to the historical tonal classes or these syllables/morphemes.

4.1.7. **Present analysis** Instead of deriving melodies on
polysyllabic expressions from syllable tones (either
synchronic citation tones or historical tonal classes of
individual syllables), the present analysis takes the
position that there are a set of tonal melodies in the
language, which are lexically associated with polysyllabic
and monosyllabic melody-bearing units alike.

4.1.7.1. Inventory of melodies. There are six melodies,
which are given below:

1 2 3 4 5 6
/1/ /42 1/ /42 24/ /3/ /24 5/ /5/

4.1.7.2. Monosyllabic forms. Under the present analysis,
monosyllabic citation forms are melody-bearing units on a
par with polysyllabic ones, with no particular status such
as the basic form or the underlying form. The contours on
monosyllables are the same as those on polysyllabic forms
but for one thing. When the above melodies are associated
with monosyllabic forms, there is a restriction against
the occurrence of melodies with initial 42. The melodies
that occur on monosyllables are as follows:

1 2 3 4
/1/ /3/ /24/ /5/
/4/
The /4/ variant of the third melody occurs only on syllables with the glottal ending.

Although the restriction against the occurrence of certain melodies in monosyllables seems arbitrary in the present analysis, alternative analyses deriving longer forms from monosyllables face the more serious problem of accounting for the occurrence of /42/ in only polysyllabic forms. As Lyu's chart (which predicts disyllabic melodies from historical tonal classes of the two syllables) shows, /42/ can occur on any kind of syllable and in the environment of any kind of syllable. It is therefore difficult to treat /42/ as a variant of any of the citation melodies.

4.2. Domain of tones in Danyang

In the last section, we argued that melodies in polysyllabic expressions in Danyang cannot be taken as a function of the tones of the component syllables either from the point of view of citation tones or historical tonal classes. It is now necessary to address the question of what the melody-bearing unit is in Danyang.

Although many examples of melody-bearing items from previous discussion happen to be words, it seems doubtful, though, that the melody-bearing unit is the grammatical word. Lyu (1980) specifically warns that the domain of tonal changes is not 'word' but 'zi zu (character/syllable
group). He gives da feng 'big wind' and hao ren 'good person' as examples which are zi zu but not words.

4.2.1. Phonological units as domain

The particular examples given by Lyu, i.e. da feng and hao ren seem to suggest that the melody-bearing unit can be larger than the word. Although this is an argument that the melody-bearing unit in Danyang and the grammatical word are not coextensive, it still does not rule out the possibility that the melodies are assigned to syntactic units like the phrase. To establish the phonological status of the melody-bearing unit, we have to show that it does not correspond to any morphosyntactic unit.

Given Lyu's argument that the melody-bearing unit can be larger than the word, we will now show that the melody-bearing unit can be smaller than the terminal elements on syntactic trees. In compounds, a melody can be assigned to each part. For example, in the following compound two melodies, each enclosed in square brackets, occur:

\([42 \ 1] [5 \ 5]\)
bo li cha bei

glass tea cup

Since two melodies are found on one syntactic word, the syntactic word cannot be the same as the melody-
bearing unit. It might be thought at this point that although the melody-bearing unit does not correspond to the terminal nodes of the syntactic tree, it does correspond to the parts of the compound which are words too. To show that this is not true, we give the following example:

\[
\begin{align*}
de & yi & zhi \\
'Germany'
\end{align*}
\]

The transliteration of 'Germany' is a simple word consisting of only one morpheme. The melodies of [1 1] and [24] are found on the morphemic disyllabic and monosyllabic sequences respectively. So in fact, the argument for the phonological status of melody-bearing unit in Danyang is even stronger. The melody-bearing unit can correspond to non-morphological unit as well. The melody-bearing unit cannot be anything but a phonological unit of some sort.

Having established the phonological status of the melody-bearing unit, let us now turn to the question of what kind of phonological unit it is. We concluded earlier that Danyang is not a syllable tone language like Mandarin, since tonal contours of longer expressions are not a function of those on the component syllables. So the phonological unit cannot be the syllable. It has to be a larger phonological unit.
4.2.2. Prosodic structure in Danyang

Further evidence for the phonological nature of the melody-bearing unit in Danyang comes from the non-isomorphy between morphosyntactic structures and melody-bearing prosodic structures in trisyllabic and quadrisyllabic forms. The non-isomorphy is seen most clearly in the fact that different morphosyntactic structures may give rise to the same prosodic structure. More specifically, left- and right-branching morphosyntactic structures can both have left-branching prosodic structures. This shows that prosodic constituents need not be morphosyntactic constituents.

An examination of the mapping between morphosyntactic and prosodic structures in Danyang reveals that there is a tendency to maintain and develop left-branching structures in general. This tendency is seen in three ways: 1) the default prosodic structure in morphosyntactically simple trisyllabic forms is \([x \ x] \ x\); 2) left-branching morphosyntactic structures keep their configurations in prosodic structures; 3) right-branching morphosyntactic structures give rise to left-branching prosodic structures but not vice versa.

In addition to favoring left-branching structures, we also see a preference for \([x \ x][x \ x]\) structures in quadrisyllabic forms. This preference is realized in two
ways: 1) the default prosodic structures in morphologically simple quadrisyllabic forms is \([x \ x][x \ x]\); 2) morphosyntactic structures of the configuration of \([x \ x][x \ x]\) keep their configuration in the morphosyntax-phonology mapping.

To see all this, let us look at trisyllabic forms and quadrisyllabic forms in turn. In the following exposition, the prosodic structures given are induced from Lyu's description of the distribution of tonal patterns on expressions of different lengths. When a single melody is distributed on two, three or four syllables, the syllables seem to form one prosodic constituent with no internal structure.

4.2.2.1. Trisyllabic forms. In trisyllabic forms, the possible morphosyntactic structures are:

\[
\begin{align*}
&[ [ x \ x ] \ x ] \\
&[ x \ [ x \ x ] ] \\
&[ x \ x \ x ]
\end{align*}
\]

In other words, we have left-branching, right-branching and no branching. The corresponding prosodic structures, judging from the distribution of melodies that are found on forms with such syntactic configurations, are not necessarily the same however.
When the structure is flat, the resulting prosodic structure can be left-branching, i.e. the first two syllables are grouped together and assigned one of the six disyllabic melodies while the last syllable has its own melody. Some of the examples from Lyu are given below:

\[
\text{[42 \ 1] \ [3]} \\
gong \ nong \ bing \\
'worker, peasant, soldier'
\]

\[
\text{[3 \ 3] \ [1]} \\
zuo \ zhong \ you \\
'left, center, right'
\]

\[
\text{[24 \ 5] \ [24]} \\
shang \ zhong \ xia \\
'up, middle, down'
\]

\[
\text{[42 \ 24] \ [24]} \\
si \ wu \ liu \\
'four, five, six'
\]

As expected, flat structures also take trisyllabic melodies:

\[
\text{[ 1 \ 1 \ 1 ]} \\
fa \ lan \ xi \\
'France'
\]

When the structure is left-branching, i.e. the first two syllables form a morphosyntactic constituent, the prosodic structure is left-branching as well. There are three varieties. The first variety is formed by simply attaching the third syllable bearing a separate melody to a disyllabic prosodic constituent. The melodies on both parts are undisturbed:
The second kind is where a melody-bearing prosodic constituent is formed from the first two syllables only when the third syllable is attached. Examples of this variety are as follows:

\[
[3 \ 3 \ ] [1] \\
huo che zhan \\
fire car station \\
'rail way station'
\]

\[
[42 \ 24 \ ] [24] \\
tai ping yang \\
supreme peace ocean \\
'the Pacific Ocean'
\]

\[
\text{-----} \rightarrow [24] [24] [24] \\
\text{\quad} [42 \ 24] [24] \\
\text{\quad} li fa dian \\
\text{\quad} \text{trim hair shop} \\
\text{\quad} \text{'barber shop'}
\]

\[
\text{-----} \rightarrow [3] [3] [3] \\
\text{\quad} [24 \ 5] [3] \\
\text{\quad} liu sheng ji \\
\text{\quad} \text{keep sound machine} \\
\text{\quad} \text{'phonograph'}
\]

The third variety is rare, as pointed out by Lyu. When the third syllable with its melody is added, the disyllable changes its melody. An example of this variety is given below:

\[
\text{-----} \rightarrow [3 \ 3 \ ] [3] \\
\text{\quad} [42 \ 24 \ ] [3] \\
\text{\quad} mo li hua \\
\text{\quad} \text{jasmine flower}
\]
In addition to giving rise to left-branching prosodic structures, left-branching morphosyntactic structures occasionally also have single trisyllabic melodies, as shown below:

\[
\begin{align*}
\text{shanghai ren} & \quad \text{[1 1 1]} \\
\text{'Shanghai person'} & \\
\text{he zuo she} & \quad \text{[1 1 1]} \\
\text{'co-op shop'} & \\
\text{nian shu ren} & \quad \text{[1 1 1]} \\
\text{read book person} & \\
\text{'student'} & 
\end{align*}
\]

In sum, then, flat structures and left-branching morphosyntactic structures give rise to left-branching prosodic structures as well as allowing trisyllabic melodies. This contrasts sharply with right-branching structures, where in addition to giving rise to right-branching prosodic structures and trisyllabic melodies we also find left-branching prosodic structures.

In most cases, trisyllabic melodies are used:

\[
\begin{align*}
\text{xi tie si} & \quad \text{[3 3 3]} \\
\text{slender iron wire} & \\
\text{zhen hong mu} & \quad \text{[42 1 1]} \\
\text{real red wood} &
\end{align*}
\]

But right-branching structures can also give rise to left-
branching prosodic structure:

\[
\begin{array}{c}
[1] [42 \ 24] \\
----> [42 \ 24 ] [1] \\
da \ ma \ lu \\
big \ horse \ path \\
'boulevard'
\end{array}
\]

\[
\begin{array}{c}
[3] [42 \ 1] \\
----> [42 \ 24 ] [3] \\
tie \ gong \ ji \\
iron \ male \ chicken \\
'miser'
\end{array}
\]

As a contrast, flat and left-branching morphosyntactic structures never give rise to right-branching prosodic structures.

Lyu mentions that when a monosyllable is combined with a disyllable, the structure remains right-branching if it is an expression with verb-object relationship between the parts:

\[
\begin{array}{c}
[24] [3 \ 3] \\
kan \ xiao \ shuo \\
read \ novel
\end{array}
\]

But this fact should be taken with caution in describing the prosodic structures of the language, as almost all the other examples he gives are nominal expressions consisting of heads and modifiers. Hence there will be no basis for comparison if we include the v-o combinations. We will then exclude this from further consideration.

Summarizing the above in chart form, we find the
relationship between morphosyntactic and prosodic structures as follows: (The '!' mark indicates that the pattern is preferred. The question mark shows the doubtful relevance of the right-branching expressions with v-o components, in view of the fact that almost all the other examples are modifier-head nominal expressions)

### PROSODIC

<table>
<thead>
<tr>
<th></th>
<th>LEFT</th>
<th>FLAT</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>morph</td>
<td>LEFT</td>
<td>!x</td>
<td>x</td>
</tr>
<tr>
<td>morph</td>
<td>FLAT</td>
<td>!x</td>
<td>x</td>
</tr>
<tr>
<td>synt</td>
<td>RIGHT</td>
<td>x</td>
<td>!x</td>
</tr>
</tbody>
</table>

The above shows:

1) The language favors left-branching prosodic structures. For flat morphosyntactic structures, to the extent that there is prosodic constituency, it is left-branching;

2) And to the extent that there is mismatch between morphosyntactic structures and prosodic structures, it is the right-branching structures that take on left-branching prosodic structure rather than vice versa;

3) Whereas right-branching prosodic structures, if
possible at all, can result only from right-branching morphosyntactic structures, left-branching structures can result from configurations of all kinds.

We can thus conclude that in compounds the default prosodic structure is left-branching.

4.2.2.2. Prosodic structures in names and transliterations. The default left-branching structure is also seen in personal names and transliterations of foreign words. A few words about Chinese names are in order here. Chinese names have two parts, the surname and the given name. Most Chinese surnames are monosyllabic, the exceptions being ouyang, sima, situ and a few others. The majority of given names are disyllabic. So the majority of Chinese names are trisyllabic, with the surname preceding the given name. My name can serve as an illustration:

SURNAME GIVEN NAME
Zhang Zheng-sheng

Surnames do not have any meaning, even if the characters for them happen to denote morphemes otherwise. Given names vary as to their meaningfulness. People tend to stay away from transparent semantic constituents, although the parts of the given name may be desirable characters in their meaning. So if there is any structure to
It is interesting however to note that the phonological constituency in trisyllabic names are different. Take Mandarin for example. As is the case with trisyllabic compounds and phrases, the phonological constituency for trisyllabic forms is \([x x x]\), based on evidence from the discussion of the fast speech lenition rule and the tonal assimilation rule in 3.3. The same phonological organization seems to be true in the case of personal names in Mandarin.

Chao (1968, p.37) reports that 'in rapid speech, weakened initials tend to become voiced continuants; retroflex \(j\), \(ch\), \(sh\), all becoming \(r\), and palatal \(j\), \(ch\), \(sh\) all becoming \(i,...\)'. When my name is spoken fast, the second syllable tends to be slurred and only the retroflex component and the high central vowel are left. It also leans on the preceding syllable:

\[
\text{Zhang Zheng-sheng} \\
[\text{dzraig} \text{drii} \text{rii}] \rightarrow [\text{dzraigri} \text{rii}]
\]

The Second-to-First Tone Rule also applies in names. One of my friends' name can be pronounced in normal speed and fast speech respectively as:

\[
35 35 35 \rightarrow 35 55 35 \\
\text{wangliangguo} \quad \text{wangliangguo}
\]
As we reformulated it in 3.3, the rule applies in a trisyllabic form only when the first two syllables form a prosodic constituent.

The same left-branching structure is found, when names are pronounced in Danyang. An example of a personal name in Danyang is given below:

\[
\begin{array}{c}
1 & 1 & [4] \\
bo & ju & yi \\
\end{array}
\]

'name of famous Tang poet'

If the name has any internal structure at all, it will be right-branching, as the first syllable is the surname and the last two syllables are the given name.

In transliterations, left-branching prosodic structures are also observed:

\[
\begin{array}{c}
24 & 5 & [3] \\
ni & gu & ding \\
\end{array}
\]

'Nicotine'

\[
\begin{array}{c}
5 & 5 & [24] \\
jia & na & da \\
\end{array}
\]

'Canada'

\[
\begin{array}{c}
42 & 24 & [3] \\
na & po & lun \\
\end{array}
\]

'Napoleon'

Left-branching prosodic structure is also observed in half transliterations, where some syllables reflect the meaning of the original word as well as the sound. When the last
syllable is meaningful, the resulting left-branching prosodic structure is not surprising:

\[
[3 \ 3] [5]
\]
ai ding bao
ai ding castle
'Edinburgh'

What is worth noticing is when the first syllable is meaningful and yet we still get left-branching structure:

\[
[5 \ 5] [3]
\]
bing qi ling
ice qi iing
'ice cream'

4.2.2.3. Quadrisyllabic forms. The possible morphosyntactic structures in quadrisyllabic forms are:

\[
\begin{align*}
[x \ x \ x \ x]
[x \ x] [x \ x]
[[x \ x \ x] \ x]
[[[x \ x] \ x] \ x]
[[x [x \ x]] \ x]
[x [x \ x \ x]]
[x [[x \ x] \ x]]
[x [x [x \ x]]]
\end{align*}
\]

When the morphosyntactic structure is flat, the preferred prosodic pattern is [[x x][x x]], although a single, quadrisyllabic melody is used as well:
When the structure is \([[x \ x][x \ x]]\), the preferred pattern is still \([[x \ x][x \ x]]\), the quadrisyllabic melody being possible too:

\[
[42 \ 24] [5 \ 5]
\]
\[
she hui ke xue
'social science'
\]
\[
[3 \ 3 \ 3 \ 3]
bei jing da xue
'Beijing University'
\]

When the structure is \([[x \ x \ x] x]\), the prosodic structure tends to be isomorphic with the morphosyntactic structure:

\[
[5 \ 5 \ 3 ] [24]
gao er fu qiu
'golf ball'
\]

When the structure is \([[[[x x] x] x]]\), i.e. left-branching all the way, the prosodic structure depends on what the prosodic structure of the trisyllabic part is in isolation. When it has left-branching internal structure, it is \([[x x] x x]\):

\[
[42 \ 24] [3] [24]
jiao ta che hang
foot peddle car shop
'bike shop'
But when the trisyllabic component has a trisyllabic melody, the prosodic structure is \([x\ x\ x]\ x\) or it bears a quadrisyllabic melody. There is no specific examples of this kind in Lyu, but as he takes all \([x\ x\ x]\) trisyllabic forms to behave the same way when a syllable is added it will not make any difference whether the trisyllabic form has left-branching or right-branching structure. Thus, the example with original \([[x\ [x\ x]]\ x]]\) as morphosyntactic structure, presented next, can be used as well.

When the structure is \([[x\ [x\ x]]\ x]]\), the prosodic structure also depends on what the prosodic structure of the trisyllabic form is in isolation. When it has a trisyllabic melody, the prosodic structure is \([[x\ x\ x]\ x]]\) or bears a quadrisyllabic melody:

\[
\begin{array}{llll}
[42 & 1 & 1 & 1][24] \\
qing\ gong\ ye\ ju \\
'light\ industry\ bureau'
\end{array}
\]

\[
\begin{array}{llll}
[42 & 1 & 1 & 1] \\
qing\ gong\ ye\ ju \\
'light\ industry\ bureau'
\end{array}
\]

But when the trisyllabic form has a right-branching internal structure, the prosodic structure is \([[x\ [x\ x]]]\ x]]\):

...
When the structure is \([x \ [x \ x \ x]]\), it has a quadrisyllabic melody:

\[
\begin{array}{llll}
42 & 1 & 1 & 1 \\
xin & ying & ge & lan \\
'New England'
\end{array}
\]

When the structure is \([x \ [x \ [x \ x]]]\), that is, right-branching all the way, it bears a quadrisyllabic melody as well:

\[
\begin{array}{llll}
42 & 1 & 1 & 1 \\
zhen & jin & jie & zhi \\
real & gold & ring
\end{array}
\]

When the structure is \([x \ [[x \ x] \ x]]\), the prosodic structures of \([[x \ x] \ x\] or the quadrisyllabic melody are used:

\[
\begin{array}{llllll}
3 & 3 & 3 & 3 & 1 \\
laohuo & che & zhan & old & fire & car \\
station & 'old & train & station'
\end{array}
\]

\[
\begin{array}{llllll}
3 & 3 & 3 & 3 & 3 \\
laohuo & che & zhan & old & fire & car \\
station & 'old & train & station'
\end{array}
\]

The correspondence between morphosyntactic structures and prosodic structures can be summarized as follows (The '
')' indicates morphosyntactic brackets; the ' [ ] ' indicates melody-bearing prosodic units; the '! ' shows that the pattern is preferred; and the '?' shows that there are questions about the structure shown:

<table>
<thead>
<tr>
<th>MORPHOSYNTACTIC</th>
<th>FLAT</th>
<th>LEFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>F ( x x x x )</td>
<td>![[x x][x x]]</td>
<td>[x x x x]</td>
</tr>
<tr>
<td>L (x x)(x x)</td>
<td>![[x x][x x]]</td>
<td>[x x x x]</td>
</tr>
<tr>
<td>T</td>
<td>![[x x][x x]]</td>
<td>[x x x x]</td>
</tr>
<tr>
<td>((x x x) x)</td>
<td>[x x x x]</td>
<td>![[x x x][x x]]</td>
</tr>
<tr>
<td>L ([(x x)x]x)</td>
<td>[x x x x]</td>
<td>?[[x x][x x][x]]</td>
</tr>
<tr>
<td>E ([(x x)x]x)</td>
<td>[x x x x]</td>
<td>![[x x x][x x]]</td>
</tr>
<tr>
<td>F ([(x x)]x)</td>
<td>[x x x x]</td>
<td>?[[x][x x][x x]]</td>
</tr>
<tr>
<td>T ([(x x x)]x)</td>
<td>[x x x x]</td>
<td>![[x x x][x x]]</td>
</tr>
<tr>
<td>R (x (x x x))</td>
<td>![[x x x x]]</td>
<td>[x x x x]</td>
</tr>
<tr>
<td>G (x((x x)x))</td>
<td>![[x x x x]]</td>
<td>[x x x x]</td>
</tr>
<tr>
<td>H (x(x(x x)))</td>
<td>![[x x x x]]</td>
<td>[x x x x]</td>
</tr>
</tbody>
</table>

In left-branching morphosyntactic structures above, square brackets are used together with parentheses. They indicate the prosodic structure of the trisyllabic form before it enters into the quadrisyllabic form. Such intermediate prosodic structure seems to matter. The prosodic structures with '?' are undetermined with respect to their branching. They result from left-branching morphosyntactic structures. So they are grouped with left-branching prosodic structures. They may not be
analysed as such in the final analysis.

The pattern can be summarized as follows:

1) When the morphosyntactic structure is flat or \([x\ x][x\ x]\), the prosodic pattern of \([x\ x][x\ x]\) is preferred, although the quadrisyllabic melody is also used;

2) When the morphosyntactic structure is left-branching, i.e. \([x\ x\ x]\), \([[[x\ x]\ x]\ x]\) and \([x\ [x\ x]\ x]\), the prosodic structure is in general isomorphic with the morphosyntactic structure. But the trisyllabic part can have a single trisyllabic melody even when it has internal morphosyntactic structure. When such is the case, a single quadrisyllabic melody can also be used.

3) When the pattern is right-branching, i.e. \([x\ [x\ x]\ x]\), \([x\ [x\ [x\ x]]]\) and \([x\ [[x\ x]\ x]]\), a quadrisyllabic melody is used; although for the last structure a left-branching \([x\ x\ x]\ x]\) is used in addition as well.

As we can see, there are only two things that have to be said specifically about the prosodic structures of quadrisyllabic forms; one is about the construction of the structure \([x\ x][x\ x]\), when the morphosyntactic structure is flat; and the other is about the restructuring involved in right-branching cases, which gives either a quadrisyllabic melody or a left-branching prosodic structure. Taking all the data together, we conclude that the preferred patterns here are \([x\ x][x\ x]\), left-
branching structures in general, and the quadrisyllabic melody.

4.2.2.4. Stress Patterns

In addition to tonal information, Lyu also gives some information on the stress patterns in Danyang. Interpreting it with our terminology, we can summarize it as: within a melody-bearing unit, the stress is left-dominant; but between melody-bearing units, the stress is right-dominant. This can be illustrated with the following example:

```
/ \  \\
  / \  \\
 s  w  s
[5 5] [24]
```

But as Lyu assumes that the melody on the first two syllables arise from tonal sandhi and the last syllable exhibits citation tone, he characterizes the situation by saying that within a sandhi tone group the first syllable is strong and the citation tone is strong. Lyu does not talk about the strength relationship between the weaker members in trisyllabic and quadrisyllabic 'sandhi tone group', i.e. melody-bearing unit in our terminology.
4.2.3. *Trisyllabic melodies and quadrisyllabic melodies*

4.2.3.1. The possibility of hierarchical structures.

Forms with trisyllabic melodies and quadrisyllabic ones have been treated as having no internal prosodic structure. There are a number of reasons to suspect that these flat structures are left-branching in origin as well. Let us look at trisyllabic melodies and quadrisyllabic ones in turn.

Distributionally, trisyllabic, flat melodies co-occur with left-branching structures, as we can see below:

<table>
<thead>
<tr>
<th></th>
<th>LEFT</th>
<th>FLAT</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MORPHOSYN</td>
<td>LEFT</td>
<td>!x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>FLAT</td>
<td>!x</td>
<td>x</td>
</tr>
<tr>
<td>TAX</td>
<td>RIGHT</td>
<td>x</td>
<td>!x</td>
</tr>
</tbody>
</table>

In fact, the flat and left-branching structures almost exhaust the possible prosodic structures. If we want to make a unified generalization, flat and left-branching ones should form a natural class rather than a disjunction. Since the language favors left-branching
prosodic structures, we should expect that structures bearing trisyllabic melodies also have left-branching prosodic structures.

Tonally, flat structures differ from left-branching structures only in the last syllable. Whereas the last syllable has an independent melody in left-branching structures, in flat structures it has no tone and gets its tone by spreading from the left. The following two examples exemplify the difference:

<table>
<thead>
<tr>
<th>LEFT-BRANCHING</th>
<th>FLAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>'dynasties'</td>
<td>sheep hair shirt</td>
</tr>
</tbody>
</table>

In the first case, we have two melodies; but in the second case, there is only one.

We can therefore hypothesize that flat structures developed from left-branching ones. The process can perhaps be understood as paralleling the development of trochaic forms from iambic forms in Mandarin. Iambic forms in Mandarin have full tones on both syllables but trochaic forms only have tone on the first syllable. The condition for the development seems to be idiomatization and in general trochaic forms are more opaque semantically than iambic compounds. In Danyang, flat structures seem to be used in very colloquial contexts, as Lyu pointed
out.

We still have to explain the fact that when right-branching structures are restructured, the preferred pattern is flat rather than left-branching. The explanation seems to be that the flat structures are preferred because of the need to keep the forms in question distinct from originally left-branching structures.

The quadrisyllabic melody can analogously be analysed. Let us look at its distribution again (The '(' indicates morphosyntactic brackets; the '[' ]' indicates melody-bearing prosodic units; the '!' shows that the pattern is preferred and '?' shows that the structure is in doubt):

<table>
<thead>
<tr>
<th>MORPHOSYNTACTIC</th>
<th>PROSODY</th>
</tr>
</thead>
<tbody>
<tr>
<td>F ( x x x x )</td>
<td>FLAT</td>
</tr>
<tr>
<td>L (x x) (x x)</td>
<td>![[x x][x x]]</td>
</tr>
<tr>
<td>T (x(x x))</td>
<td>![[x x][x x]]</td>
</tr>
<tr>
<td>(x x) x</td>
<td>![[x x][x x]]</td>
</tr>
<tr>
<td>L (x x) x x x</td>
<td>![[x x][x x][x]]</td>
</tr>
<tr>
<td>E (x x) x x x</td>
<td>![[x x][x x][x]]</td>
</tr>
<tr>
<td>F (x(x x)) x</td>
<td>![[x x][x x][x]]</td>
</tr>
<tr>
<td>T (x(x x)) x</td>
<td>![[x x][x x][x]]</td>
</tr>
<tr>
<td>R (x (x x))</td>
<td>![[x x][x x][x]]</td>
</tr>
<tr>
<td>G (x((x x)x))</td>
<td>![[x x][x x][x]]</td>
</tr>
<tr>
<td>H (x(x(x x)))</td>
<td>![[x x][x x][x]]</td>
</tr>
<tr>
<td>T (x((x x)x))</td>
<td>![[x x][x x][x]]</td>
</tr>
</tbody>
</table>
The quadrisyllabic melody serves as the less preferred pattern except where right-branching structures are restructured. Since some quadrisyllabic melodies are the less preferred variants of left-branching structures, we may relate the two the same way as in the case of trisyllabic forms. The development of flat quadrisyllabic melodies from left-branching structures can be similarly explained as in the case of trisyllabic forms, i.e. as a concomitant of the 'idiomatization' process paralleling the development from iambic to trochaic forms in Mandarin. The fact that flat structures are used as the preferred, rather than the variant, patterns for right-branching morphosyntactic structures can perhaps be explained the same way, that is, in terms of maintaining the distinction from originally left-branching structures.

The possibility of hierarchical structures in trisyllabic and quadrisyllabic melodies needs more thorough investigation.

4.2.4. Foot as the domain of melodies

We have seen that the melody-bearing unit in Danyang is not the syllable. The possibility of more than one melodies in simple words indicates that the melody-bearing unit is smaller than the word. The fact that the preferred prosodic structures of the language seem to be built from disyllabic and monosyllabic units strongly
suggests that the melody-bearing unit is the foot in Danyang.

In addition to disyllabic and monosyllabic feet, Danyang also has trisyllabic and quadrisyllabic feet, although to a limited degree. Depending on whether or not trisyllabic and quadrisyllabic forms have internal structure, feet in Danyang would be either binary and recursive or n-ary and non-recursive.

4.3. Foot in Mandarin and Danyang

When we compare Danyang and Mandarin, we see quite a bit of similarity in having the prosodic unit foot and its construction.

The unit of foot is relevant for Mandarin, as we see in chapter 3. In Mandarin, the foot is used as the domain of the Third Tone Rule and probably the Second-to-First Tone Rule as well.

The similarity between Danyang and Mandarin does not stop at the fact that both languages employ the same unit. The tendency to favor left-branching prosodic structures is also found in Mandarin. The default structure is left-branching as seen in the application of the Third Tone Rule in transliterations like somali 'Somali':

```
so ma li
3 3 3
[ ]f by DM
2 3 3 cycle 1
2 2 3 cycle 2
```
Cyclic application, which is independently argued for in chapter 3, cannot give the correct result unless the prosodic structure is left-branching.

We argued in the context of Mandarin too that in fast speech right-branching structures become left-branching ones in order to account for the fast speech pronunciation of mai hao jiu 'buy good wine', which can be derived by assuming the same structure as above for Somali:

```
mai hao jiu
buy good wine
3 3 3
[  ]f by DM
2 3 3 cycle 1
2 2 3 cycle 2
```

Such restructuring is supported by the application of the Second-to-First Tone Rule and segmental lenition under fast speech context. Both of these rules require the foot in a left-branching structure as the domain of application as well. In fact, the output of the last example can further undergo the Second-to-First Tone rule:

```
mai hao jiu
2 2 3
2 1 3
```

A question that naturally arises at this point is
whether the prosodic unit foot and the tendency to favor left-branching structures is also found in other Chinese dialects. As Danyang lies between Mandarin dialects and Wu dialects, its similarity with Mandarin is not too surprising. Is the prosodic structure of Danyang influenced by Mandarin? If so, do other Mandarin dialects exhibit the same prosodic structures? Do other language families show the same kind of prosodic structure? All this awaits future research.

4.4. The role of disyllables

The special status of disyllabic forms in Danyang is better understood in the larger context of Chinese. Lyu (1963) discusses the tendency for Chinese to favor disyllabic combinations. Only a few phenomena that Lyu mentions will be used as illustration here.

First, monosyllabic words are giving away to disyllabic ones in modern Chinese. Many disyllabic compounds are formed whose parts are identical in meaning and hence are semantically redundant. A few examples from Lyu are given below:

```
shen    ti    wei    da    tian    di    yi    fu
 body    body    great    big    field    earth    clothes    clothes
 'body'    'great'    'field'    'clothes'
```
On the other hand, trisyllabic words are shortened into disyllabic ones. For example:

\[
\begin{align*}
\text{luo hua sheng} & \quad \longrightarrow \quad \text{hua sheng} \\
\text{drop flower birth} & \quad \text{flower birth}
\end{align*}
\]

'peanut'

\[
\begin{align*}
\text{wai guo yu} & \quad \longrightarrow \quad \text{wai yu} \\
\text{out country language} & \quad \text{out language}
\end{align*}
\]

'foreign language'

\[
\begin{align*}
\text{long jing cha} & \quad \longrightarrow \quad \text{long jing} \\
\text{dragon well tea} & \quad \text{dragon well}
\end{align*}
\]

'Longjing tea'

Another manifestation of the tendency to favor disyllabic expressions can be seen in the fact that, disyllabic abbreviations are most commonly found in the language. Some examples are given below:

\[
\begin{align*}
\text{Beijing daxue} & \quad \longrightarrow \quad \text{Beida} \\
\text{Beijing university}
\end{align*}
\]

\[
\begin{align*}
\text{Qinghua daxue} & \quad \longrightarrow \quad \text{Qinghua} \\
\text{Qinghua university}
\end{align*}
\]

\[
\begin{align*}
\text{shehuizhuyi jiaoyu} & \quad \longrightarrow \quad \text{shejiao} \\
\text{socialist education (a movement in the 60's)}
\end{align*}
\]

\[
\begin{align*}
\text{tudi gaige} & \quad \longrightarrow \quad \text{tugai} \\
\text{land reform}
\end{align*}
\]

The third way in which disyllables are favored can be seen indirectly in the prevalent use of quadrisyllabic set phrases and idioms in Chinese. Many of these have the internal structure of \((x\ x)(x\ x)\). Two examples are given below:
Chen (1979) argues that poetic scansion for Chinese classical regulated verse proceeds by the Foot Formation Rule, i.e. by grouping syllables into disyllabic feet, subject to constraints imposed by morphosyntactic structures. Shih (1986) argues that a variation on Chen's Foot Formation algorithm can define the basic domain of the Third Tone Sandhi Rule as well.

Outside of Chinese, we find Lehiste's comment (1970:163):

The disyllabic sequence, consisting of an odd and even numbered syllable, appears as a basic phonological building block out of which words seem to be constructed. Indeed, the rules for the occurrence of quantity contrasts in Finnish and Estonian are most economically formulated not just in terms of syllables but in terms of disyllabic sequences.
Notes to Chapter 4

1. For a discussion of the syllable types in Chinese, the reader is referred to chapter 5, in the section on Shanghai citation tones.

2. The distinction between colloquial and reading pronunciations in Danyang is recognized in Lyu (1980). It may ultimately explain much of the variation observed in Danyang tones. The details of the colloquial/reading dichotomy are not clear to me from Lyu's description, however.

3. This was brought to my attention by Mary Beckman.

4. Lyu (1980) takes some tonal pattern for a particular combination of syllables to be primary while the variant patterns to be secondary. The fact still remains that the secondary patterns have to be accounted for in addition to the primary patterns.

5. We will then be describing the prosodic structures in a restricted area only. But since our purpose here is to contrast the different behaviors of left- and right-branching structures in the mapping between morphosyntax and prosodic structures, the restriction is necessary.

6. The question is whether or not the structures should be binary-branching, with the first two constituents grouped together first. In other words, should we have $[[x \ x] \ [x]] \ [x]$ and $[[[x] \ [x \ x]] \ [x]]$, or the flatter structures $[[x \ x] \ [x] \ [x]]$ and $[[x] \ [x \ x] \ [x]]$, given in the chart? Following the morphosyntactic structures, we should have the more hierarchical structures; but the evidence from tonal melodies does not motivate but is merely consistent with more structure.
CHAPTER V

WORD TONES IN SHANGHAI

As a contrast to Mandarin and Danyang, some Wu languages such as Shanghai should best be described by taking the word as the unit of tonal representation underlyingly. This chapter provides explicit arguments for doing so.

The word tone characteristics of Wu languages did not go unnoticed. George Kennedy (1953:373) comments, in discussing the two tone patterns in Tangsic, a Wu language, that 'One might be led to the broad guess that the function of tone in Tangsic is essentially not to distinguish otherwise homophonous syllables, but to express syntactic relationships'. Sherard (1972) suggests the idea of 'tonal envelop' on phonological words in Shanghai. He remarked that (116) '...polysyllabic words are learned as discrete entities associated with a particular overall tone contour, and not as combinations of individually identifiable syllables'. In the recent dissertation by Shih (1986), languages in the Wu family are classified as word tone languages distinct from syllable tone languages such as Mandarin and Cantonese.
But in these previous works, arguments for the word tone status of Wu languages were either lacking or insufficient.

Let us first examine the different criteria for word tones that have been proposed or may be proposed. After seeing the inadequacy of these criteria, we will propose a criterion against which Shanghai will be measured.

5.1. What is a word tone?

In the tonal literature, there are a number of ways in which the distinction between word tones and other kinds of tones is characterized. But before we look at these criteria, we should keep in mind a distinction that is often overlooked, i.e. the distinction between morpheme tones and word tones.

5.1.1. Morpheme Tones versus Word Tones

A distinction between syllable tone languages and word tone languages is made for Chinese languages in Shih (1986). While the classification may in fact be a correct one, it is important to point out that there are two differences between syllable tones and word tones and not one. If languages are without morphology and the morpheme is coterminous with the word, a simple distinction between syllable tones and word tones is adequate. But all
languages, including Chinese, have some morphology, at least compounding. We should then make a distinction between morpheme tones and word tones, at least conceptually.

In fact, the issue of whether to represent tones on the morpheme or the word is not a trivial one. Analysts have differed on exactly this issue. Edmondson and Bendor-Samuel (1966), for example, claim that the tonal patterns on words in Etung are best regarded as a feature of the phonological word, whereas Welmers (1962) argues that in the analysis of Kpelle, tone should be expressed as a feature on morphemes. In fact, Welmers makes a stronger claim in another context (1973:80):

There seems to be no known language in which pitch is significant only for units larger than a morpheme (such as words), but smaller than a phrase.

Given the non-trivial distinction between morpheme tones and word tones, we should entertain the following picture, with the distinction between phonological and morphological tones on the one hand and between morpheme tones and word tones on the other:

UNITS OF TONAL ASSIGNMENT

/ \ phonological morphological
/ \ syllabic morphemic word-level
A complicating factor in Chinese is that the morpheme is often taken as coterminous with the syllable and hence something like the following is assumed:

```
UNITS OF TONAL ASSIGNMENT
/ \  
syllabic  word-level
morphemic
```

There are two considerations that should be kept in mind when we adopt the diagram above. First of all, although the morpheme can be coterminous with the syllable, they are still distinct notions and should be kept strictly apart on a theoretical level. Secondly, it is always possible that even within Chinese, languages may differ as to the degree of overlap between the morpheme and the syllable.

5.1.2. Tone-bearing unit

The unit of tonal representation sometimes is defined in terms of tone-bearing units, that is, the tone-bearing unit in a word tone language is the word but in a syllable tone language it is the syllable. But, tone-bearing unit, seems a confusing piece of terminology, since the syllable can also bear tones in what are characterized as word tone languages. The term is at least used in two different senses in the field. In one sense, tonal melodies are said to be associated with tone-
bearing units, typically, vowels (Yip 1980); this sense of tone-bearing unit is crucially different from the sense in which words are said to be the tone-bearing unit in word tone languages.

The different senses of the term tone-bearing unit stem from the different senses of the term tone. Apart from the distinction between phonetic pitch and phonological tone, there is a third sense of tone implicitly assumed in calling the word the tone-bearing unit, that is, tones as tonal melodies, or tonal patterns. These tones are not to be confused with units such as H and L, variously called tonemes, tonal features or tonal autosegments. Using tone-bearing units to distinguish between word tones and other kinds of tones makes sense only in this third sense of tone. Although the term tone-bearing unit can be used with a specific designated meaning, it nonetheless is a bad piece of terminology and causes no end of confusion. In the dissertation, I choose to use the term melody-bearing unit for the third sense of tone and use pitch-bearing unit to refer to the segmental material to which tones are associated. The term tone-bearing unit is used only when no possibility of confusion arises.

But even if we take care in using terminology, using melody-bearing units to define word tone is still less than satisfactory. It is simply stating that word tones
are properties of the word, namely, word tones! And syllable tones are properties of the syllable, namely, syllable tones!

5.1.3. Melody-counting

More informative is to identify word tones, as opposed to syllable tones, by looking at the combinatorial possibility of tones occurring in multisyllabic combinations. The criterion goes like this: syllable tone languages give rise to $T^n$ melodies in a word, where $T$ is the number of citation tones in the language and $n$ the number of syllables in the combinations in question. Word tone languages have considerably fewer melodies, the number being only $T$, regardless of the number of syllables in the words. Informative and rigorous as it seems, melody-counting is also problematic in a number of ways.

The melody-counting criterion can conceal differences between languages that are all said to have word tones. There are at least two types of word tone languages, depending on whether the language has underlying word tones or simply an underlying syllable tone system with drastic neutralization at the word level. If melody counting is done on the surface forms of both types of languages, there will be no difference between the two. On the other hand, if the counting were done before the neutralization rules are applied, we would get
174

an entirely different result. But as far as I can see, there is no practical utility in using the criterion when we already know what the underlying form is. For then we will surely know where tones are represented.

In the surface kind of word tone languages, melody-counting at the surface level measures the degree of neutralization; but neutralization is neither necessary nor sufficient for the existence of word tones. So it can be misleading in identifying word tones.

As it is formulated, the melody-counting criterion cannot be used to differentiate between morpheme tones and word tones. When it is normally used to differentiate syllable tones and word tones, it is actually used in setting apart syllable tones from morpheme tones. In complex words with more than one morpheme, the possible number of melodies will not stay constant regardless of the number of syllables. It would of course be possible to extend the use of the criterion and say that in word tone languages the number of word melodies will stay constant regardless of how many morphemes there are in the word. But I am not aware of such a usage.

5.1.4. Citation Tones as Syllable Tones

In addition to the criteria mentioned above for identifying word tones, there is often an implicit criterion used to identify syllable tones in Chinese. We
will show that it is mistaken too.

In the Chinese tonal literature, the reference to 'citation tone' is frequently found. The theoretically neutral interpretation of citation tone is that it is the tone that occurs on syllables uttered in isolation\(^3\), as opposed to tones that occur on syllables in combination with other syllables. Since citation tones are nothing but tones that happen to be uttered on monosyllables, it is entirely conceivable that they can be morpheme tones or even word tones. Citation forms, always meaningful, should more sensibly be interpreted as monosyllabic words rather than syllables.

There are, however, two unfortunate assumptions that are often associated with reference to citation tones. One is that the citation tone is the underlying form from which sandhi forms are derived in phonological environments. This assumption has been challenged by a number of researchers (Hashimoto 1982, Pan 1986, Chan 1986) on the basis of data from a number of Chinese languages showing distinctions found in sandhi forms not present in citation forms.

The second hidden assumption often attached to citation tones is that they are properties of the syllable that carry them. The mistaken assumption perhaps arose from the fact that monosyllabic minimal pairs are found in the language. The existence of monosyllabic minimal pairs
easily leads to the conclusion that the language is a lexical tone language in the sense of Pike (1948), one with contrastive pitch on every syllable. It also accounts for the desire to derive longer melodies from citation tones.

As pointed out above, citation tones can be word tones or morpheme tones that happen to be associated with monosyllables. Minimal pairs in citation forms do not necessarily show that the language has contrastive pitch on every syllable. For all we know, the melodies that show up on monosyllabic words might not occur on syllables in other contexts. We will see that in Shanghai exactly this situation is found.

Instead of deriving word melodies from citation tones, it is entirely possible to map melodies that occur on combination forms onto citation monosyllables. Taking citation tones to be word tones can explain two phenomena difficult to account for in analyses that derive word tones from citation tones by spreading. First, the tonal contours of the whole word might not be the same as those of the citation tones. This is first observed by Sherard (1972:92) for Shanghai. Under a word tone analysis, this is rather to be expected, because whole melodies might not fit onto monosyllables and hence are partly unassociated. The second difficulty for spreading analyses is seen in situations where there are more melodies than citation
tones in the language. Danyang serves as an example of this, as we saw in the last chapter. A word tone analysis can handle this with restriction on the occurrence of some word melodies on monosyllables.

To conclude, the existence of citation tones -- even the existence of minimal pairs in citation forms does not establish the language as having syllable tones.

5.1.5. Cyclic versus Post-cyclic

To remedy the shortcoming of the melody-counting criterion, which does not distinguish morpheme tones as opposed to word tones, we may use as our criterion the level at which tones are associated with pitch-bearing units.

In Pulleyblank's (1983) incorporation of tonal phonology into the lexical phonology framework, a distinction was made between cyclic and non-cyclic assignment of tones. Since the cycle corresponds to morphological strata and the non-cyclic level often corresponds to the end of morphological operations, it may be tempting to conceptualize the distinction between morpheme tones and word tones as the difference between cyclic assignment and non-cyclic assignment in the lexical phonology framework.

While morpheme tones can be assigned cyclically,
non-cyclic assignment does not have to be restricted to assigning word tones. There are two ways in which non-cyclic assignment can map morpheme tones onto words.

The first situation can be demonstrated with Clements & Ford's (1979) analysis of Kikuyu. In Kikuyu, we find the following concatenation of morphemes:

\[ \text{mo} + \text{e} + \text{rEk} + \text{ang} + \text{er} + \text{i} + \text{E} \]
\[ \text{L} \quad \text{H} \quad \text{L} \quad \text{H} \]

What is of interest here is that the morpheme tones and the segmental materials remain unassociated at the end of the concatenation process. The association then is non-cyclic:

\[ \text{mo} + \text{e} + \text{rEk} + \text{ang} + \text{er} + \text{i} + \text{E} \]
\[ \text{L} \quad \text{H} \quad \text{L} \quad \text{H} \]

(In Kikuyu, the first tone is associated with the second morpheme by an initial association rule, and the mapping of the rest of the tones is achieved by Clements and Ford's version of the Well-formedness Condition.)

In the situation above, the number of melodies is not reduced and hence the word tone is a simple composition of the morpheme tones. Such a language still should be considered a morpheme tone language.

We now turn to the second way in which morpheme
tones can be mapped non-cyclically. The situation can be illustrated hypothetically here (Shanghai has been analyzed in exactly this fashion by Zee & Maddieson (1979):

\[
\begin{array}{cccc}
M_1 & M_2 & M_3 & M_4 \\
[ ] & [ ] & [ ] & [ ] \\
T_1 & T_2 & T_3 & T_4
\end{array}
\]

In the process of morphological concatenation, the tones of all but the first morpheme are deleted:

\[
\begin{array}{cccc}
M_1 + M_2 + M_3 + M_4 \\
[ ] \\
T_1
\end{array}
\]

Non-cyclical assignment then proceeds, in whatever fashion the language-specific conventions dictate. Although there is reduction of melodies and the tone is assigned at the word level, the language will still be a morpheme tone language, because the morphemes all have underlying tones and the tone assigned post-cyclically is that of the first morpheme.

5.1.6. Word Tones are Lexical Word Melodies

It is now time to put forward our own criterion for word tones. It is given as follows:

A word tone is a word melody that is lexically represented on the whole lexical item, be it a simple word or a complex one. It cannot
be derived from either the tones of the component morphemes or those of the component syllables.

In other words, the relationship between word tones and those on component morphemes and syllables is not compositional. The term 'non-compositional' is reminiscent of the relationship often found between word meanings and component morpheme meanings. As Aronoff (1976) argues, the minimal sign should be the word rather than the morpheme; words as minimal signs can take on meanings deviant from, or even totally unrelated to, the meanings of their morphemes. There is no reason why words cannot also take on properties of sounds, tones or otherwise, that are deviant from or unrelated to their components.

It is possible for languages to have word tones in addition to morpheme tones, since it is possible for some words to acquire word tones but for other words to have tones that are derivable through regular processes. But a word tone language will have predominantly word tones.

In word tone languages, tonal melodies remain constant not only regardless of the number of syllables but also without regard to the number of morphemes.

It will be seen in Shanghai that there are word melodies unaccounted for by deriving them from tones of morphemes or syllables; and it is advantageous to analyze
the language as a word tone language.

5.1.7. Some Word Tone Languages. Despite Welmers' (1973) claim to the contrary, languages have been analyzed as having the word as the domain of tonal features. Etung, as Edmondson and Bendor-Samuel (1966) analyzed it, is a word tone language. Mazaudon (1973) argues that Tamang, a language spoken in Nepal, should also be analyzed as a word tone language. In Shih (1986), Lhasa Tibetan and a number of Wu languages are analyzed as having word tones.

A largely pre-theoretical description of Lopa, a language spoken in the southeastern part of Tibet, by Sun et al. (1980) may let us infer that it is a word tone language as well. A quote from Sun et al. is given below in translation (119):

The first syllable of every word is mid level, the one following is high level; short vowels occurring in the last position of a polysyllabic word are read as falling. For example, in ogo 'five', ici 'water', dugu 'honey' the second syllable should be read as falling. In addition to the above, tones vary with context; for example, ta: 'listen' in isolation is read as high level, but ta:pong 'heard' should read ta:pong [mid-high]. Or api: 'four' reads a pi: [mid-high], ani 'two' reads as a ni [mid-fall]; but the word formed from these two words 'eight' pi:ni reads as pi:ni [mid-fall], pi: becoming mid level. Also, porok 'chicken' reads as porok [mid-fall], rokne 'hen' reads rokne [mid-fall]; the two roks are pronounced with different tones depending on the position.
We may infer from the above that Lopa has the word tone of mid-high, which comes out as mid-falling when the final syllable is short. Sun et al. comment that there are no contrastive tones in Lopa but that a word or a syllable normally has a certain pitch. All this baffling state of affairs results from looking at the syllable as the relevant unit of tonal representation.

5.2. Word Tones in Shanghai

In this section, we argue that tones should be represented on the word rather than on the syllable or the morpheme in Shanghai. Three arguments will be given. The first argument is based on the distribution of contour tones in polysyllabic and monosyllabic morphemes. It shows that if Shanghai is treated as a syllable tone language, there would be severe restriction on the distribution of contour tones. The second argument shows that the spreading analysis used in most previous studies runs into problems in deriving word melodies by spreading initial tones rightward. The third argument presents word melodies that are unaccounted for by deriving them from the component morphemes or syllables.

But before we look at the arguments, some background on the language and its tones are in order.

Shanghai belongs to the Wu language family, one of several large language families in China. It is spoken in
the metropolis Shanghai in the Southeastern province of Jiangsu. Of the many languages in China, Shanghai is one of the most studied. A detailed description of the tone sandhi facts in Shanghai can be found in Xu et al. (1981, 1982, 1983) and Shen (1981). Generative treatments of Shanghai phonology include Walton (1971), Sherard (1972), Zee & Maddieson (1979), Yip (1980), Wright (1983), Jin (1985) and Selkirk & Shen (1988). Spectrographic studies of Shanghai tones can be found in Z&M.

5.2.1. Tones in Shanghai

5.2.1.1. Citation "Tones". In Shanghai every monosyllable uttered in isolation belongs to one of five types. Following are the five types of syllables, with their segmental make-up, pitch contours and example morphemes ([h] is voiced aspiration or murmur; [?] is a glottal stop and N stands for a nasal segment):

<table>
<thead>
<tr>
<th>Types</th>
<th>structure</th>
<th>contour</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td>C V (N)</td>
<td>53</td>
<td>[pi] 'edge'</td>
</tr>
<tr>
<td>B.</td>
<td>C V (N)</td>
<td>24</td>
<td>[pi] 'change'</td>
</tr>
<tr>
<td>C.</td>
<td>Ch V (N)</td>
<td>13</td>
<td>[phi] 'skin'</td>
</tr>
<tr>
<td>D.</td>
<td>C V ?</td>
<td>5</td>
<td>[pi?] 'pen'</td>
</tr>
<tr>
<td>E.</td>
<td>Ch V ?</td>
<td>13</td>
<td>[phi?] 'other'</td>
</tr>
</tbody>
</table>

The pitch contours of these tones have been given in the pretheoretical notation used in Chinese tonology. The five types of syllables have four distinct tonal contours,
the C and E types both having 13. Despite the fact that only four tonal contours are found, the five types of syllables are said to be five tones, that is, different sheng 'voice/sound', in traditional Chinese tonology.

The term sheng 'voice/sound' in traditional Chinese phonology has generally been taken to mean 'tone' in the phonological literature. It is important to point out, however, that 'tone' is understood differently in traditional Chinese phonology from its conception in western linguistic tradition, because different 'tones' in the Chinese traditional sense not only contrast in their pitch contours but also in such things as the voice/phonation of the initial consonant and the presence or absence of stops at the end. Sherard (1972) chose to call tones in Chinese 'syllable types' instead.

Stripping away the purely tonal contrasts between the syllable types in Chinese, we still find at least the following possible contrasts among syllable types that arise non-tonally:

```
(A) $ (B) $ (C) $ (D) $ 
C V X Ch V X C V K Ch V K
```

The C/Ch contrast in the initial position involves either difference in voicing or phonation type (depending on different dialects); the X/K contrast in the final position gives rise to a distinction between unchecked and
checked syllables. The unchecked syllables end with X, where X can be nasals or glides or nothing at all. The checked syllables end with K, where K is from the set of stops \{p, t, k, ?\}. Essentially then, there are two things to be manipulated, initials and final consonants, and four combinations are possible.

Not all Chinese languages have all four types of syllable structures. Mandarin, for example, has neither checked syllables nor syllables with voiced/murmured initials. Cantonese has checked syllables but not voice/murmured initials. The Wu dialects have both murmured initials and checked syllables.

The C and E types in Shanghai are different in initial/phonation from the rest. For Z&M, the difference is a matter of voicing in the initial obstruents, the C and E types having voiced obstruent initials. For them, sonorants pattern with voiceless obstruents in being the initials of only A, B, D types. But for Xu et al.(1981, 1982, 1983), a Chinese source, the difference is in phonation types, the C and E types having murmured or voiced aspiration. Xu et al.'s inventory of consonants also has aspirated/murmured sonorants and unaspirated/unmurmured sonorants as contrastive sets. The D and E types are also different from the other types because they end with a glottal stop. Syllables starting with voiced initials or murmured phonation all tend to
have lower pitch than their unvoiced or unmurmured counterparts; and syllables with glottal stop endings all tend to be short. Z&M's study of the duration of syllables in isolation shows the following result:

- A. 262 msec
- B. 327 msec
- C. 352 msec
- D. 102 msec
- E. 147 msec

As we can see, D and E are shorter than the others, due to the glottal stop endings. And C and E, which have murmured phonation, are longer than their unmurmured counterparts. E actually shows the interaction between phonation and glottal stop ending rather nicely, since it is intermediate in length between the long and the short syllables.

5.2.1.2. "Tones" on longer forms. Let us give the tonal contours of multisyllables up to four syllables long. To facilitate the comparison between them and those of monosyllables, the two kinds of contours are presented side by side. The phonetic values of tonal contours are extracted from Z&M's tracings of fundamental frequencies, with a few minor differences of interpretation from theirs (T' means 'raised tone').
The phonetic values are also presented in Z&M's fashion, i.e. by specifying a tonal value for every syllable. It should be noted that another way of transcribing the contours has been used in the literature as well. In Sherard (1972), for example, contours are transcribed by using lines that more iconically represent the change in pitch. Shih (1986) adopts the same representation. Two examples are given below, illustrating the use of the notation:

\[ \text{fi ji} \quad \text{'aircraft'} \quad \text{pa? pa?} \quad \text{'800'} \]

As can be seen, there is an alternative pattern for the quadrisyllabic forms. This fact will be important later on, when we argue against analyses that spread tones of initial syllables rightward in accounting for melodies of polysyllabic combinations.

5.2.2. Tonal Melodies in Shanghai Polysyllabic morphemes

Unlike Mandarin, which has few polysyllabic
morphemes in the native vocabulary, Shanghai has quite a number of polysyllabic morphemes that cannot be conceivably compounds. In fact, having polysyllabic morphemes seems one of the features that distinguish Wu dialects in general from Mandarin dialects. In his article, 'The monosyllabic myth', George Kennedy argues cogently against the view that Chinese is a monosyllabic language, citing examples from Tangsíc, his native language in the Wu family.

In Shanghai, polysyllabic morphemes are quite easy to find. Some examples are given below:

\[
\begin{array}{ccc}
H & L & M \\
\text{pi?se} & \text{tɕikong} & \text{tɕiokue} \\
'rascal' & 'very' & 'very'
\end{array}
\]

The polysyllabic morphemes are also found in the grammatical morphemes of Shanghai. In the pronoun system, we find the following:

\[
\begin{array}{ccc}
L & H \\
\text{a?la?} & \text{yila?} \\
'I, we' & 'they'
\end{array}
\]

Among the aspectual particles, we find:

\[
\begin{array}{c}
L & H \\
\text{leʔhE} \\
'progressive'
\end{array}
\]

The existence of polysyllabic morphemes is not in
itself an argument for word tones in Shanghai. After all, Mandarin has polysyllabic morphemes as well, only to a smaller extent and mainly in transliterations of foreign words. The argument for word tones in Shanghai can be seen in the difference between Mandarin and Shanghai in the distribution of contour tones. In Mandarin, contour tones are distributed no differently in polysyllabic morphemes from compounds. Every syllable in Mandarin polysyllabic morphemes can have any of the contour tones found in monosyllabic morphemes, at any position:

\[
\text{LH LH LH LH LH LH} \\
[\text{pu tao ya}_m \text{ hai}_m \text{ mei}_m \text{ wan}_m]
\]

'still not finished'

'Portugal' 'not yet finished'

Indeed, this fact is used in chapter 2 as an argument for the syllable tones in Mandarin. But in Shanghai, we find a different situation. On monosyllabic words, we generally find contour tones only. (The exception is syllables with glottal ending, which tend to be rather short and abrupt.) But in polysyllabic monomorphemic words, we do not in general find contour tones on individual syllables. (The exception to this general rule is in morphemes with murmured phonation in the first syllable and glottal stop in the final syllable, where contours rise at the end of the syllable.) The contrast between the melodies of monosyllabic morphemes and those
of polysyllabic morphemes is illustrated with the following:

<table>
<thead>
<tr>
<th>monosyllabic morphemes</th>
<th>polysyllabic morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>LH Ngo</td>
<td>LH yi la?</td>
</tr>
<tr>
<td>'I'</td>
<td>'you'</td>
</tr>
<tr>
<td>HL Ti</td>
<td>HL tsiokue</td>
</tr>
<tr>
<td>'sky'</td>
<td>'book'</td>
</tr>
</tbody>
</table>

If we adopt the syllable tone analysis for Mandarin, we might say that there are four tonemes in the examples given so far, L, H, LH and HL. These are associated in the lexicon with syllables, as they are argued to be in Mandarin. Then we find severe restriction on the occurrence of tones on syllables. In polysyllabic morphemes, out of the sixteen possible combinations in disyllabic sequences, only two, [H.L] and [L.H], are found, both of them consisting of level tones. In monosyllabic morphemes, only two melodies, [LH] and [HL], are found out of four logical possibilities, both of them contour tones.

A more reasonable analysis would be to assume two tonemes, LH and HL. These two melodies are represented on the morpheme and mapped onto the syllables by association conventions. Level tones are found when the number of syllables is equal or greater than that of tones, and
contour tones result when the number of syllables is smaller than that of tones. The results of the mapping for the examples above are as follows:

<table>
<thead>
<tr>
<th>monosyllabic morphemes</th>
<th>polysyllabic morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>LH</td>
<td>LH</td>
</tr>
<tr>
<td>.writeFile</td>
<td>L H L H</td>
</tr>
<tr>
<td>ngO</td>
<td>yi la?</td>
</tr>
<tr>
<td>'I'</td>
<td>'they'</td>
</tr>
<tr>
<td>ti</td>
<td>tsio kue</td>
</tr>
<tr>
<td>'sky'</td>
<td>'very'</td>
</tr>
</tbody>
</table>

Such an analysis, which may seem alien to the syllable-tone oriented point of view, is motivated linguistically. George Kennedy observed in his "Two tones patterns in Tangsic" (1953:373):

'It is beyond question that tone is word-distinctive on monosyllables in Tangsic. But these monosyllables form a very small percentage of the vocabulary....From spoken material alone it is extraordinarily difficult, as I have often discovered, to make up a word list for Tangsic in terms of syllables with fixed tones. There is a limited number of nouns and verbs that can be given tones in isolated form; but for the large part of the vocabulary only an appeal to a hypothetical system of writing can settle the question of tone. It has always been hard for sinologists not to feel that the written graphs are in fact the final authority. In a realistic treatment of Tangsic they cannot be '.

It is interesting how people still manage to give underlying tones for every syllable. One wonders how this
is done. To start with, these submorphemic syllables always occur in combination and are never able to bear the attributed citation tones. There are a few ways these citation tones could have been contrived. First, there is a practice, when transcribing dialects, to use the same characters to represent homophones, due to the lack of characters for submorphemic syllables. The citation tones of the segmentally homophomous morphemes could have been assigned to these syllables. Second, etymological and historical tones may have been assigned to these syllables. Thirdly, the tones on the initial syllables of these simple morphemes could have been circularly derived from the tones of the whole morpheme, by applying the spreading rule backward—a 'squeezing rule'.

5.2.3. Relatedness of Melodies

The relatedness between melodies of the polysyllabic forms and those of monosyllabic forms is quite easy to see when we put them side by side:

<table>
<thead>
<tr>
<th>type &amp; tone of 1st syl.</th>
<th>disyllabic combination</th>
<th>trisyllabic combination</th>
<th>quadrisyllabic combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. HL</td>
<td>H L</td>
<td>H M L</td>
<td>H M L' L</td>
</tr>
<tr>
<td>B. MH</td>
<td>M M'</td>
<td>M H L</td>
<td>M H M L</td>
</tr>
<tr>
<td>C. LH</td>
<td>L M'</td>
<td>L H L</td>
<td>L H M L</td>
</tr>
<tr>
<td>D. H</td>
<td>M' H</td>
<td>M' H L</td>
<td>M' H M L</td>
</tr>
<tr>
<td>E. LH</td>
<td>L M'</td>
<td>L L M'</td>
<td>L L L LM'</td>
</tr>
</tbody>
</table>

(L H M L)
It is an unmistakable fact that in polysyllabic forms the number of tonal contours remains constant regardless of the number of syllables. Furthermore, the number is far lower than would have resulted from a simple concatenation of monosyllabic citation tones. It is also quite apparent that in general levels tones occur on multisyllabic forms while contour tones occur on monosyllables.

The discovery of the relatedness can be attributed to Sherard (1972), who said:

'The tone contour over the entire word can be thought of as being dependent on the syllable type of the initial syllable, S1, although the resultant contour does not necessarily have the same shape as that of the tone of the initial syllables....' (92)

The mere relatedness of the melodies is consistent with three positions:

1) tones are represented on the syllable; in polysyllabic combinations, all but the tone of the first syllables are deleted and the tone of the first syllable spread rightward to cover the whole combination.

2) tones are represented on the morpheme; in compounds, all but the tone of the first morpheme are deleted and the tone of the first morpheme spread rightward to cover the entire word.

3) tones are represented on the word, on polysyllables and monosyllables alike.
5.2.4. Against Spreading

5.2.4.1. Spreading Analyses. Due to the strong resemblance between word melodies and those of the initial syllables in isolation, most analyses (Zee & Maddieson 1979 being the most explicit) assume deletion of non-initial tones and subsequent spreading of the initial tonal melodies to envelop the whole word. The Z&M analysis is given below. The derivation of most of the disyllabic forms is achieved by the following rules in conjunction with the WFC:

Word Boundary deletion.

When a compound is formed, the internal word boundaries are deleted.

Tone Deletion 1.

\[
\begin{align*}
\text{#S}_1+\text{S}_2+\text{S}_3+\text{S}_4+\ldots\# & \quad \text{---->} \quad #\text{S}_1+\text{S}_2+\text{S}_3+\text{S}_4+\ldots\#\\
\text{|| | | | | | | |} & \quad \text{|| | | | | | | |} \\
\text{#T}_1 \text{ T}_2 \text{ T}_3 \text{ T}_4 \ldots\# & \quad \text{#T}_1 0 0 0 \ldots\#
\end{align*}
\]

Condition: does not apply if \( T_1=\text{MH}/\text{LH} \) and \( T_2=\text{H} \) and \( S_3\ldots=0 \)

The effect of this rule is to delete all the tones of the non-initial syllables except when the second syllable is of the D type and the initial syllable is A. The exceptional clause is to account for the contours of the AD type and should not concern us here.

Association rule (a).

Delete the association line between the second of two tones and \( S_1 \) in any compound.
Following the WFC, this rule completes the right-spreading.

A sample derivation of the compound di ti 'escalator' is given below:

a. Word boundary deletion applies.

b. Tone deletion applies:

\[
\begin{array}{c}
\text{#} \text{d}i + t\text{i}\# \\
\hline
|\text{LH} & \text{HL}\#
\end{array} \\
\begin{array}{c}
\hline
\text{#} \text{d}i + t\text{i}\#
\end{array}
\]

c. WFC applies:

\[
\begin{array}{c}
\text{#} \text{d}i + t\text{i}\# \\
\hline
|\text{LH} & 0\#
\end{array} \\
\begin{array}{c}
\hline
\text{#} \text{LH}\#
\end{array}
\]

d. Association rule (a) applies:

\[
\begin{array}{c}
\text{#} \text{d}i + t\text{i}\# \\
\hline
|\text{LH} & #
\end{array} \\
\begin{array}{c}
\hline
\text{#} \text{LH}\#
\end{array}
\]

After word boundary deletion applies, tone deletion of all but the initial syllable tones occurs; the WFC spreads the tones of the second part of the first syllable automatically rightward. The last part delinks the original association between the H part of the melody and the first syllable.

As we can see, Zee & Maddieson's analysis is spreading in its most literal form. The H part of the melody first spreads and then gets delinked from its
original syllable. The inelegance of Z&M's analysis partly results from their decision to link the tones and the syllable before compounding.

Less literal spreading analyses like those of Yip (1980), Wright (1983), and Selkirk & Shen (1988) map the melodies of the initial syllable onto the whole word after those of the non-initial syllables are deleted. There is no delinking of tones from the original syllable since the association is rather late, probably post-cyclical. Although in these theories, tones of the first syllable/morpheme are not linked early on to their segmental materials but mapped onto the word directly, they are still not word tones lexically.

The difficulties for both kinds of spreading analysis are apparent when we try to derive explicitly tonal contours of polysyllabic forms from monosyllabic forms.

5.2.4.2. The case of EX type. Words starting with E type initial syllables and any type of non-initial syllables are the most problematic for the spreading analysis. E type syllables start with breathy phonation and end with a glottal stop. Its tonal contour is characterized by a final rise on the last syllable of the word, regardless of the number of syllables in the word.
It is hard to see how an underlying representation of the initial syllable can be posited that will make the spreading analysis work.

Three possibilities for dealing with the situation suggest themselves. First, we may use right-to-left association and spreading for words of this kind. Alternatively, we may entertain 'edge in' association, which specifies the tones for the edges of the word and then fill in the middle portion. Both of these analyses assume not only language-particular but syllable type particular association conventions and are therefore quite ad hoc.

The third option is to use lexical pre-linking, which links the LM to the last syllable in the lexicon. But that means the unit of lexical representation cannot be anything but the word. In other words, a word tone analysis will be assumed.

5.2.4.3. Polysyllabic morphemes

In 5.2.2., polysyllabic morphemes are used as evidence that tones are not properties of the syllable but of the morpheme or the word. The argument is based on the lack of contour tones on these morphemes.

These morphemes can also be used as an argument.
against the spreading analysis. As the initial syllables in these morphemes do not occur in isolation, they cannot be demonstrated to have their own lexical tones. There are then no initial tones to serve as the basis of spreading.

5.2.5. Lexical Word Tones.

The strongest evidence of word tone comes from alternative tone patterns that are found on words with otherwise identical component tones. There are two ways in which the alternative patterns are realized. They may be used in addition to the regular patterns; or they may replace the regular patterns.

5.2.5.1. Alternative tone patterns in B, C types

In Shanghai (Xu et al. 1981, 1982, 1983), apart from the regular patterns of word melodies, we also find alternative word melodies peculiar to particular lexical items:

<table>
<thead>
<tr>
<th>Citation Tone</th>
<th>Regular Pattern</th>
<th>Alternative Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;old&quot;</td>
<td>L H lo si</td>
<td>H L lo si</td>
</tr>
<tr>
<td>&quot;teacher&quot;</td>
<td>L H lo si</td>
<td></td>
</tr>
<tr>
<td>&quot;pants&quot;</td>
<td>M H ku tsi</td>
<td>H L ku tsi</td>
</tr>
<tr>
<td></td>
<td>M H</td>
<td></td>
</tr>
</tbody>
</table>
In the two examples above, alternative patterns are used in addition to regular patterns. But in the following, only the alternative patterns are allowed:

\[
\begin{array}{ll}
*L & H \\
H & L \\
kun & kao \\
'sleep' & \\
\end{array}
\quad \begin{array}{ll}
*M & H \\
H & L \\
si & ka \\
'world' & \\
\end{array}
\]

In trisyllabic forms, alternative patterns are also observed:

<table>
<thead>
<tr>
<th>citation tone</th>
<th>regular pattern</th>
<th>alternative pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>of 1st syl.</td>
<td>LH</td>
<td>(\ddot{\text{n}}) jiu sang</td>
</tr>
<tr>
<td></td>
<td>(\ddot{\text{n}}) jiu sang</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\ddot{\text{m}}) jiu sang</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'study'</td>
<td>'graduate student'</td>
</tr>
<tr>
<td></td>
<td>M H L</td>
<td>ku jia? kuo</td>
</tr>
<tr>
<td></td>
<td>ku jia? kuo</td>
<td>ku jia? kuo</td>
</tr>
<tr>
<td></td>
<td>'pants'</td>
<td>'pants sleeves'</td>
</tr>
<tr>
<td></td>
<td>(\ddot{\text{m}}) jiu sang</td>
<td></td>
</tr>
</tbody>
</table>

If the regular pattern is derived from the tones of the citation tone of the first syllable by spreading, according to the most prevalent analyses of Shanghai, then it is hard to see how the alternative patterns can also be derived in the same way. The alternative pattern seems rather to be assigned to the lexical items as a whole at the word level.

These examples show that the spreading rule, if it exists, would not be regular. There are two possible ways to cope with this in the spreading analysis. One way
would be to say that the tonal melodies on these words are exceptionally marked in the lexicon. But we wonder why exceptional marking assigns only the melody of HL and not other conceivable ones, like HH, LL, or H, L. The second way is to posit an alternative underlying tonal melody, H L, for these words. This way, we explain why only HL, and not any other melody is obtained. But an alternative underlying tonal melody is just as unjustified. The morphemes bearing the alternative tonal melody are for all purposes the same morphemes occurring elsewhere with different tonal melodies. The alternative melody can never be obtained as the citation forms of the first syllables of these morphemes. The postulation of an alternative underlying tonal melody in the cases at hand is undesirable both for its abstractness and the lack of independent motivation.

5.2.5.2. Alternative pattern for quadrisyllabic E type words. We have been glossing over a detail in the presentation of the pattern for E type words, a detail that is seldom noticed but critically important. In Xu et al.'s data set, we find that for quadrisyllabic words having E type initial syllables, there is an alternative pattern. The main and the alternative patterns are listed below:
The alternative pattern is identical to the pattern for C type words. Suppose we adhere to the spreading analysis. We would again be committed to two underlying tonal contours for the same E citation tone and again be burdened with problems outlined above for such solutions. But in this case, the problem is more serious. Why do we have the alternative pattern for the quadrisyllabic forms only? In a word tone analysis, the alternative tonal pattern can well be associated with quadrisyllabic lexical items only.

5.2.6. The present analysis

In the previous two sections, three arguments were given against deriving melodies on polysyllabic expressions from tones of initial syllables by spreading. The first argument has to do with the difficulty of giving an underlying tone for E type syllables. The second argument depends on the existence of polysyllabic morphemes in Shanghai, whose initial syllables cannot be said to have independent melodies. The third argument is based on the alternative melodies on polysyllabic forms.

The present analysis takes the word, rather than the
syllable or the morpheme, as the lexical melody-bearing unit. In doing so, it takes seriously the suggestion by Sherard that "polysyllabic words are learned as discrete entities associated with a particular overall tone contour, and not as combinations of individually identifiable syllables."(116) It is interesting to note that while later analyses (Z & M 1979; Wright 1983) attribute spreading to Sherard, I could find no mention of spreading in his work.10.

In the following, the inventory of word tones in Shanghai and the way these word tones are realized on polysyllabic and monosyllabic words is discussed.

5.2.6.1. Three word melodies

In tonal literature published in China, Shanghai is said to have five citation tones. As we have seen, the tones (sheng 'voice') in Chinese traditional descriptions should be treated as syllable types, as the differences in such tones are not only tonal but also segmental. The actual tonal contrasts are fewer than five, as Z & M (1979) and Wright (1983) show. Previous descriptions, including Z & M and Wright, all base the inventory of tone melodies on monosyllabic citation forms, however.

In our word tone analysis, the inventory of word melodies in Shanghai can be arrived at by looking at the
melodies occurring in expressions of different lengths:

<table>
<thead>
<tr>
<th>MONOSYLLABIC</th>
<th>DISYLLABIC</th>
<th>TRISYLLABIC</th>
<th>QUADRISYLLABIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. HL</td>
<td>H L</td>
<td>H M L</td>
<td>H M L' L</td>
</tr>
<tr>
<td>B. MH</td>
<td>M M'</td>
<td>M H L</td>
<td>M H M L</td>
</tr>
<tr>
<td>C. LH</td>
<td>L M'</td>
<td>L H L</td>
<td>L H M L</td>
</tr>
<tr>
<td>D. H</td>
<td>M''H</td>
<td>M' H L</td>
<td>M' H M L</td>
</tr>
<tr>
<td>E. LH</td>
<td>L M'</td>
<td>L L M'</td>
<td>L L L LM'</td>
</tr>
</tbody>
</table>

(L H M L)

One feature immediately obvious is that there is only one peak in each melody. In A type, the peak is initial; in B, C and D types, the peak occurs on the second syllable; in E type, the peak occurs finally. According to where the peak occurs, the melodies fall into three classes.

I would like to suggest that there are only three melodies in the language. The B, C and D types actually have the same melody. The differences among them can be attributed to the different segmental make-up in the different syllable types. The segmental make-up of the five tones in Shanghai is given below (h means murmur; X means zero or nasal ending):

```
A/B $        C   $        D   $        E   $        \\
/ / / / / / / / /
C  V  X      Ch V  X      C  V  ?      Ch V  ?
```

The difference between tonal melodies of B and D types for trisyllabic forms is [M H L] versus [M' H L], which can be accounted for by the absence versus presence of the
glottal stop in the initial syllable. The difference between tonal melodies of C and B/D types is \([M/M' \ H \ L]\) versus \([L \ H \ L]\), which can be accounted for by the fact that C has initial murmur and hence lower pitch (the depressor effect of voicing and murmur on the pitch is well documented in Hyman (1973)). Thus, we can conclude that for B, C and D types the melody is \([M \ H \ L]\) on trisyllabic forms. The language then has the following melodies on trisyllabic forms:

A. \([H \ M \ L]\)  B/C/D. \([M \ H \ L]\)  E. \([L \ L \ M']\)

What is distinctive in the three melodies is the location of the peak. Thus we can say that the language has three melodies, which have the H specification initially, at second position and finally respectively.

Disyllabic forms and quadrisyllabic forms are covered by the same generalization:

<table>
<thead>
<tr>
<th></th>
<th>A.</th>
<th>B/C/D.</th>
<th>E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 SYLLABLE</td>
<td>([H \ L])</td>
<td>([M M'])</td>
<td>([L M])</td>
</tr>
<tr>
<td>3 SYLLABLES</td>
<td>([H M \ L])</td>
<td>([M H \ L])</td>
<td>([L L M'])</td>
</tr>
<tr>
<td>4 SYLLABLES</td>
<td>([H M L'L])</td>
<td>([M H L'L])</td>
<td>([L L LM])</td>
</tr>
</tbody>
</table>

5.2.6.2. Citation forms

Under the present analysis, melodies are represented on lexical items, be they polysyllabic or monosyllabic. The so-called citation tones are nothing but word melodies
realized on monosyllabic words. We pointed out earlier that citation tones should not be taken as syllable tones in all cases just because they happen to be associated with monosyllables. Instead of deriving melodies on polysyllabic forms from citation forms, we map melodies onto monosyllabic forms as well as polysyllabic forms.

A comparison of the monosyllabic forms and the disyllabic forms shows that the major difference between citation contours and those on polysyllabic forms is in D type words:

<table>
<thead>
<tr>
<th>MONOSYLLABIC</th>
<th>DISYLLABIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. HL</td>
<td>H L</td>
</tr>
<tr>
<td>B. MH</td>
<td>M M'</td>
</tr>
<tr>
<td>C. LH</td>
<td>L M'</td>
</tr>
<tr>
<td>D. H</td>
<td>M'H</td>
</tr>
<tr>
<td>E. LH</td>
<td>L M'</td>
</tr>
</tbody>
</table>

But as we know that D type syllables have the glottal stop ending and are hence short, the surfacing of level tones rather than contour tones is well understood.

The fact that word melodies can be easily mapped onto citation forms is no surprise, since spreading analyses relating citation tones and tones on longer forms have been proposed in the first place.

There are three advantages of a word tone analysis over a spreading analysis:

1) A word tone analysis solves the problem of words with E type initial syllables;
2) A word tone analysis offers a better account of alternative melodies;

3) A word tone analysis can better accommodate discrepancies between melodies on polysyllabic forms and monosyllabic forms.

5.2.7. The problem of tonal representation

The analysis given in the previous section is mainly concerned with the inventory of word melodies and the mapping of these melodies onto polysyllabic and monosyllabic words.

There are a number of ways the melodies can be represented. First, as the three melodies all have one peak, an accentual analysis is possible. In fact, the three melodies may be able to be reduced to only one. Let us look at the three melodies again:


The L in A/B/C/D can be attributed to final lowering; the L in E can be attributed to the lowering effect of murmur phonation. The result is the following melody:

[M H M]

In A, the pre-H M does not surface; in E, the post-H M
does not surface.

Given the melody, there are still a number of analyses possible. According to Goldsmith's accentual analysis (1976), the lexical representation has an asterisk '*' associated with the accented syllable/vowel; the melody is introduced in the course of derivation, with '*' aligned with the 'H' in the melody and the non-H's linked to the rest of the vowels by the WFC.

Goldsmith's accentual analysis is criticized by Pulleyblank (1983). According to Pulleyblank, the introduction of the melody violates the constraint against introducing phonological entities in Chomsky and Halle (1968). He opts for the lexical linking of H to accented vowels in place of '*' and no independent melodies are used. The non-H parts of melodies are given by default tonal specification. In the case of Shanghai, the default value will be 'M'.

So far, we have assumed a fully specified representation in the sense that every syllable has a tonal specification in terms of H, L and M. It has been pointed out to me (Mary Beckman, personal communication) that such a representation distorts the true picture of the language by multiplying phonetic entities. In the underspecified representation of Pierrehumbert and Beckman (forthcoming), the H will be linked to the initial, final and the second syllable and the non-H's will arise through
phonetic interpolation.

In the absence of specific arguments at the moment, the choice between the above proposals will not be principled. What the present analysis tries to show is that the domain of melody in Shanghai is the word. To the extent that the analysis is correct, it is consistent with all the proposals above.

5.3. Domain of Tones in Shanghai

We have been assuming that in Shanghai the melody-bearing unit is the word, without making explicit the nature of the category. Is it a grammatical unit or a phonological unit? In this section, we look at this question.

Chao (1968) uses the domain of tone sandhi as a criterion for distinguishing words from phrases in Wu dialects. Thus he said (p.147):

In the Wu dialects, compound words are recognizable from their tone sandhi, which are different within words from the tone sandhi between words.

He seems to suggest that the domain of tone sandhi is the grammatical unit, the compound word.

Sherard (1972) defines the domain of Shanghai tones as the phonological word, without providing explicit argument for his choice of phonological word rather than grammatical words. He said (p.91):
The phonological word frequently coincides with the minimal grammatical unit as well in Shanghai, however in a detailed treatment of Shanghai syntax, it would probably be necessary to make a distinction between the domains of the phonological word and the syntactic word.

In another connection, he said, however (p.156):

Rarely—if ever—does the phonological word correspond to more than a single grammatical unit. ....... In the relatively few cases where a word boundary signals a lower-level (morphological) relationship, this can usually be traced to extra-syntactic factors such as utterance length; thus in general we can say that phonological patterning in Shanghai reflects syntactic patterning, up to the point where the distinction between syntactic and morphological levels become blurred as a result of frequency of usage and/or idiomatization.

The first sentence of the above quote can be taken to mean that the phonological word is not larger than the grammatical word and the rest can be taken to mean that the phonological word is not smaller than the grammatical word either. If this is true, then the notion of phonological word is not motivated on language-internal grounds.

In order to show that it is the phonological word that is the domain of tone in Shanghai, we have to establish the non-identical extension of the phonological unit and the morphosyntactic unit.

5.3.1. **Phonological word in general** In the prosodic hierarchy of Nespor and Vogel (1986), the unit of
phonological word is defined as the unit that immediately dominates the foot and is dominated by the clitic group. As far as its relationship with the grammatical word, three possibilities are entertained:

1) the phonological word is co-extensive with the grammatical word;

2) the phonological word is no larger than the grammatical word;

3) the phonological word can be larger than the grammatical words.

Nespor and Vogel demonstrate the first possibility with languages like Greek and Latin, where the phonological word is the grammatical word. They also show that in Hungarian and Turkish the phonological words can be smaller than the terminal elements on syntactic trees. But they reject the third possibility, that the phonological word could be larger than the grammatical word, and posit the clitic group for such a unit instead.

In the tonal literature, we also encounter reference to phonological word as the domain of tones. Pike (1969), in reporting the then ongoing fieldwork conducted by the Summer Institute of Linguistics on seven Tibeto-Burman languages in Nepal, considers the phonological word or the foot as the domain of tonal melodies. But he did not provide explicit arguments for doing so. A more explicit
argument for the use of phonological word as the domain of
tonal patterns is given by Edmondson & Bendor-Samuel
(1966). In their paper, tones in Etung were argued to
have the domain of the phonological word, which is
distinct from the grammatical word. Their analysis is
based on tones in verbs. Edmondson & Bendor-Samuel gives
the following description of verbal tone patterns (p.5):

Verbs consist of a root preceded by one or
more prefixes. Verbs may be marked by two
successive tone patterns, one pattern marking
the prefixes and the other the root.

In sum, the best argument for the phonological word
as a unit distinct from the grammatical word is the non-
isomorphism between phonological words and grammatical
words, although Nespor and Vogel cautioned that
'occasional isomorphism' is consistent with a universal
prosodic hierarchy.

We will see that on language-internal grounds
Shanghai tones in fact have phonological words as the
domain. Let us now turn to the evidence.

5.3.2. Non-isomorphy with morphosyntactic units

There is evidence from Shanghai that the domain of
tones cannot be always be grammatical words, since it can
be larger than this unit. Jin (1985) seems to be the
first to point out that the domain of tone sandhi extends
beyond the word. In Jin's terms, closed-class words can
enter the tonal domains of the preceding open-class words and not vice versa. The specifics of Jin's proposal are quite problematic. But it makes us aware of tonal phenomena above the level of word; and for our purposes, Jin's proposal shows that the tonal domain in Shanghai can at least have the extension of a syntactic phrase, that is, content word + function words. Function words in Shanghai can be inherently atonal or tonal. In the first case, function words get their tones from the tones of the preceding content words; in the latter case, they first lose their inherent tones and then get tones from preceding content words.

Examples of atonal function words include the so-called particles and grammatical morphemes of various kinds. Despite Jin's assignment of tones to them, it is really hard to see how they can have intrinsic tones, since they never occur in isolation nor at the beginning of tonal domains. These lexemes are all monosyllabic. Some examples are given below:

- **Question Markers**: va, a, la, a?;
- **Aspect Markers**: le?, ku, le?(hE);
- **Possessive Marker**: ge?
- **Potential/Complement Marker**: ta?
- **Classifiers**: ke?, tsa?, peng
- **Postpositions**: o, lang, li
- **Direction particles**: le, chi, zang, o, ku, jing, tsa?
Resultative particles: za?, to, tha?, ho
Discourse particles: ge?, le

Function words with inherent tones are represented by personal pronouns of various kinds. These lexemes can be monosyllabic or disyllabic. They are given below:

```
LH   LH   LH
ngo 'I' nong 'you' yi 'he/she'
L H   L H
a?la? 'we/I'   yila? 'they'
```

They can occur either independently or subsumed under the domains of other lexemes. When they occur independently, their inherent tones surface; but when they are subsumed in the domains of other lexemes, their tones are overridden by the tonal contour of the host.

Function words contrast with content words in that the latter can never be subsumed under the domains of function morphemes. Hence, their tonal contours are inviolable. The contrast is shown in the following example. (Square brackets indicate domain of tones.)

```
VP   VP
 / \ / \
PP V PP V
 / / / /
P NP P NP
[bi][zang] [go] [bi yi] [go]
LH LH HL LH HL
compare bed tall compare him tall
'longer than bed' 'taller than him'
```
In the example, when a content word zang 'bed' follows the preposition bi 'compare', it forms a separate tonal domain; but when the function word yi 'him' follows bi, the whole prepositional phrase becomes one tonal domain.

In the above example, the tonal domain consisting of a content word and a function word corresponds to a syntactic phrase. If it is the case that the tonal domain always corresponds to a morphosyntactic constituent, then we still have not established the phonological nature of the tonal domain. In fact, however, tonal domains can correspond to non-constituents as well:

\[
\text{L H M L HL} \\
\text{[pe? ngo lia beng][si]} \\
\text{give me two copy book} \\
\text{'give me some books'}
\]

The first four syllables, which bear a quadrisyllabic melody and hence form a tonal domain, do not form a syntactic constituent.

\[
\text{L H M H M H} \\
\text{[yi la?][tang ngo][a? di]} \\
\text{they beat my brother}
\]

The third and the fourth syllables bear one disyllabic melody and hence form one tonal domain. But it is easy to see that the fourth syllable must form a syntactic constituent with the following syllables.
In general, function words encliticize leftwards and form tonal domains with the preceding materials. The tonal domains formed can well be non-constituents syntactically.

Given that the domain of tone in Shanghai is phonological, we still have to find out where the unit fits in the prosodic hierarchy, i.e. whether it is the foot or the phonological word or some other unit. We just saw that it can be larger than the syntactic word, it is necessary to see whether the tonal unit can be smaller than the word.

In the last chapter, we find that the tonal unit in Shanghai cannot be the syllable, as shown by the evidence of polysyllabic morphemes. In polysyllabic morphemes, we find only level tones on the component syllables whereas in monosyllabic morphemes, we find only contour tones. If the syllable is taken as the tonal unit, we would expect contour tones on the syllables of polysyllabic morphemes as well.

In Shanghai, there does not seem to be a tonal melody mapped onto phonological units larger than the syllable but smaller than the word11. Tonal domains are largely constructed by enclitizing onto grammatical words. Therefore, it seems appropriate that Sherard uses the term phonological word for the unit of tone in Shanghai.
Notes to Chapter 5

1. The exact status of the melody-bearing unit in Shanghai will be discussed later on in 5.3.

2. Neutralization is not sufficient, because not all languages with neutralization can be said to have word tones. Mandarin has one neutralizing Third Tone Rule, but it is very much a syllable tone language, with all its tonal combinations derivable from the syllable tones. It is also difficult to say how many neutralization rules are needed for a language to become a word tone language. Neutralization is not necessary, either, for a language to have word tones. In fact, languages with underlying word tones will have them even in the absence of neutralization rules.

3. Kratochvil (1986) holds a more dismal view of the relationship between citation forms and forms in other contexts. He said: "the nature of the relationship between isolated citation forms and normal speech tones is not known (it is possible that 'demonstration tones' are names for tones comparable to the spelling names of the letters of the alphabet), and it has never been shown how the shapes of the latter could be derived from the former." (258)

4. The parallelism between sound and meaning may not be perfect here. Sound changes have been taken to be regular for the most part and meanings are subject to more idiosyncratic changes. Be that as it may, the processes of language learning, perception and production may well give rise to irregularity in pronunciation of the sort represented in word tones.

5. There is therefore a distinction between a word tone language and a word tone.

6. As a concrete example of the use of such ad hoc association, Z&M accounts for the E forms by the following language specific rule:

   Association rule (b)
   
   Insert an association line between an initial
L and the final syllable in a bisyllabic or trisyllabic syllabic compound, where $S_1=(C)V?$. The effect of this rule can be seen in the following derivation of za se, 'thirteen', which starts with E.

The Association rule (b) applies, after the Tone deletion rule I, which gets rid of the tone(s) in the second syllable:

```
#za? + se# -----> #za? + se#
\(\triangle\) 0 #
#LH -----> #za? + se#
\(\triangle\)
WFC #LH #
```

The glottal stop of the first syllable also gets deleted in the final output.

7. There may be significance to the fact that alternative patterns are found on only words with B and C type initial syllables. But I have no ready explanation of the fact at the moment.

8. It might be suggested that the variation is due to historical source or dialectal influence. But I am not aware of such factors at the moment.

9. The shift in the peak of the H tone in the alternative pattern can be better understood if an accentual version of the word tone analysis is entertained. The long-distance action involved in the shift of the H tone resembles accent shifts in accentual languages.

10. In fact, he said:

"The tone contour over the entire word can be thought of as being dependent on the syllable type of the initial syllable, $S_1$, although the resultant contour does not necessarily have the same shape as that of the tone of the initial syllables..." (92)

Three things are worth noticing here. First, Sherard talks about dependency rather than derivation between the first syllable and later ones; second, it is the syllable type of the initial syllable, rather the tone, that determines the contour of the word; and thirdly, the contours of the words are not necessarily the same as those of the initial syllables.
11. Personal names might offer some evidence that the melody-bearing unit could be larger than the syllable but smaller than the word. In other words, the situation in Danyang could hold in Shanghai as well. In trisyllabic names with a monosyllabic surname and a disyllabic given name in that order, the last syllable has an unreduced tone while the first two syllables either have reduced tones on each syllable or have a disyllabic melody. If the latter, the first two syllables form a tone domain; but if the former, the two syllables need not form a prosodic constituent.

The crucial thing then depends on whether the first two syllables, a non-syntactic constituent, bears a disyllabic melody. Further investigation is needed here.

Foreign transliterations, however, do not seem to behave the same way as personal names. Trisyllabic forms bear a trisyllabic melody rather than have an unreduced tone on the final syllable.
6.1. From syllable to word; from right-to-left. In the preceding chapters, three Chinese languages are examined with respect to their melody-bearing units, stress patterns and prosodic structure in general. There are two general tendencies that stand out:

a) The melody-bearing units become larger as we go from Mandarin to Shanghai. In Mandarin, the melody-bearing unit is the syllable; in Shanghai, it is the phonological word; but in Danyang, the melody-bearing unit is the intermediate unit of the foot.

b) As we go from Mandarin to Shanghai, the languages change from being right-dominant to left-dominant. Left- and right-dominance can be interpreted metrically. In Mandarin, only in trochaic forms do we get left-dominant patterns:

```
TROCHAIC
/ \  
s   w
```

A consistently right-dominant structure can become less so through restructuring:
Trochaic forms were historically derived from iambic forms. Restructuring occurs only in fast speech. Mandarin therefore is primarily right-dominant, but it is becoming more left-dominant.

According to Yip (1980), Shanghai has consistently left-dominant feet:

\[
\text{LEFT}
\]

In Danyang, however, we have seen a general preference for the following stress pattern, which is right-dominant at the higher level but left-dominant at the lower level. Let us call this the mixed pattern:

\[
\text{MIXED}
\]

Danyang again is intermediate between Mandarin and Shanghai. We see the comparison better in the following:
I have given the most preferred patterns in each of these languages. The more restricted patterns in Mandarin and Danyang are given below:

The mixed pattern in Mandarin is found in fast speech, as a result of restructuring; the left-dominant pattern in Danyang is found in morphosyntactically right-branching forms. We can see that in the more restricted patterns Mandarin is like Danyang and Danyang is like Shanghai. It is worth noticing that the trend is toward more left-dominant patterns rather than vice versa.

6.2. From Right to Left Dominance. The distinction between syllable-tone and word-tone is not new (Shih 1936); nor is the distinction between left-dominance and right-dominance (A. Hashimoto 1980, Chan and Ren 1986). Although one distinction has to do with the melody-bearing
unit, the other with the stress pattern, they overlap to a large extent when used to classify Chinese languages. It seems that syllable tone languages tend to be right-dominant and the word tone languages tend to be left-dominant. What is needed is to relate the two phenomena, ultimately to explain them in the same way.

I hypothesize that the tendency to favor left-dominance is the primary motivation in the concurrent changes from right-dominance to left-dominance and from syllable-tone to word-tone. This hypothesis is based on the following two considerations.

a) Primarily left-dominant languages, such as the various northern Wu languages, have word tones. This means that when a language is thoroughly left-dominant, it will not have the syllable as the melody-bearing unit.

b) In syllable-tone languages such as Mandarin, left-dominant characteristics are also observed. The left-dominant characteristics include the possibility of right-to-left branching restructuring in fast speech, the development of trochaic forms in the lexicon and the default left-branching prosodic structure as the domain of the Third Tone Rule. This means that it is not necessary for a language to have larger-than-syllable melody-bearing units to exhibit left-dominant characteristics.

Taking the above together, we hypothesize that the
tendency towards left-dominance is the underlying cause for the development from syllable tone to word tone.

6.3. From domain of sandhi to domain of melody. Left-dominance ultimately leads to the interpretation of melodies on polysyllabic forms as word or foot melodies. But this does not have to happen directly. Left-dominance can first be exhibited in the left-branching prosodic structure serving as the domain of sandhi. The domain of sandhi can then be interpreted as the domain of melody when the sandhi form on a polysyllabic form is reinterpreted as the word or the foot melody. In Mandarin, left-dominance is mainly shown in the sandhi domain. Danyang is a step further however in making the domain of sandhi the domain of melodies.

In Mandarin, the domain for the application of tonal rules is left-branching in the absence of morphosyntactic structures. For example, the Foot Formation Rule will group the first two syllables together by the DM part of the algorithm when the IC part is not applicable:

```
$ $ $ $ $ $ $ 
suo ma li 
"Somali"
3 3 3
---------- IC not applicable
[ ] DM
[[ ] ] Super-foot formed
  2 3 3 cycle 1
  2 2 3 cycle 2
```
Such left-branching prosodic structure results even in slow speech, without the restructuring seen in fast speech and regardless of whether the stress pattern is left-dominant.

The left-branching structures in Danyang seem to have started out as domains of tone sandhi but ended up being melody-bearing units when the melody in the sandhi span was reinterpreted as non-derivable from the component syllables. In the following, Mandarin and Danyang are compared in this respect (D is the domain of tone sandhi but D' is the melody-bearing unit):

\[
\begin{array}{c}
\text{MANDARIN} \\
/ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ D \\
/ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ / \ D' \\
/ \ \ \ / \ \ \ / \ \ \ / \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ so \ ma \ li \\
/ \ \ \ / \ \ \ / \ \ \ / \ \ \ / \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ gong \ nong \ bing \\
\text{'Somali'} \ 'worker, peasant, soldier'
\end{array}
\]

so ma li gong nong bing
'Somali' 'worker, peasant, soldier'

3 3 3 [3] [3] [3]

With the transition from the domain of tone sandhi to the melody-bearing unit completed, Danyang's left-branching structure is more thoroughly established. In Mandarin, the Foot Formation Rule respects morphosyntactic structures to a greater degree than in Danyang. After all, the first step in the Foot Formation Rule is to group immediate morphosyntactic constituents into feet. If the morphosyntactic structure is right-branching, the
In Danyang, however, only left-branching structures result.

Hypothesizing the tendency to favor left-dominance simultaneously explains a number of things. In addition to explaining the existence of trochaic forms and the restructuring in fast speech, it also explains why the Foot Formation Rule scans from the left when it groups syllables into disyllabic feet. That prominence must play a part in the Foot Formation Rule can be seen in the fact that at the sentence level, the scanning starts with the syllable bearing the sentential accent.
REFERENCES


He, Bao Zhang (1988). Tone Sandhi of T4 + T4 revisited. Ohio State University ms.


Lin, Tao. (1963) Beijinghua de liandu yinbian. (Sound changes in connected speech in Beijing speech.) Beijing Daxue Xuebao (Renwen Kexue) 6.


Liu, Feng-hsi (1980) Mandarin tone sandhi: a case of
interaction between syntax and phonology. Paper presented at the summer meeting of the Linguistic Society of America, Albuquerque,


Pierrehumbert, Janet and Mary Beckman (forthcoming) Japanese Tone Structure. MIT Press.

Poteet, Steve. (1985) Tone sandhi and the phonological structure of Mandarin. ms., UCSD.


