ON THE NATURE OF MANDARIN TONE AND TONE SANDHI

Hua Lin

Bachelor of Arts, Lanzhou University, 1982
Master of Education, University of Victoria, 1987

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ACCEPTED
FACULTY OF GRADUATE STUDIES

DEAN
DATE 1 Feb 93

in the Department
of
Linguistics

We accept this dissertation as conforming
to the required standard

Dr. Barry F. Carlson, Supervisor (Department of Linguistics)

Dr. Thomas E. Hukari, Departmental Member (Department of Linguistics)

Dr. James Arthurs, Departmental Member (Department of Linguistics)

Dr. Daniel J. Bryant, Outside Member (Department of Pacific and Asian Studies)

Dr. William S-Y. Wang, External Examiner (University of California at Berkeley)

UNIVERSITY OF VICTORIA

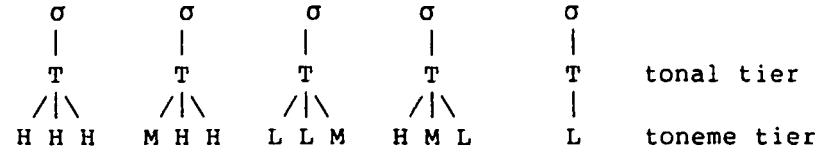
1992

Supervisor: Barry F. Carlson

ABSTRACT

Traditional representations of Mandarin tones have provided valuable descriptions of Mandarin tone sandhi processes. However, none of them has been able to to associate these processes in a principled way, or to explain why they occur at all. In contrast, I have proposed in this dissertation a unified analysis of Mandarin tones and tone sandhi, with an emphasis on the revelation of the nature of these tones and processes. Specifically, I have found that Mandarin tones are most optimally represented as follows:

(0.1)



Under such a representation, all Mandarin tone sandhi processes (i.e. the second, third, and fourth tone sandhi processes, and the neutral tone sandhi process that has been uncovered in this study) can be uniformly accounted for by the following Tone Reduction Principle:

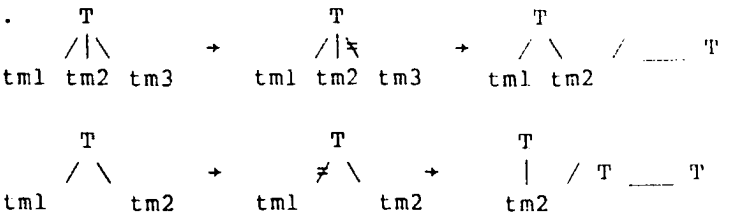
(0.2). Tone Reduction Principle

Clause A:
In normal speech, reduce a tone by one toneme iff it is immediately followed by another tone within the same prosodic foot.

Clause B:
In fast speech, reduce a tone by one toneme iff it is immediately preceded by another tone, and at the same time immediately followed by another tone within the same prosodic foot.

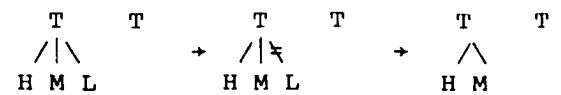
This Tone Reduction Principle functions to shorten a tone in duration by the following implementation rules:

(0.3).

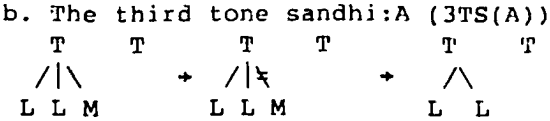


With these two rules, Mandarin tone sandhi processes can be described by the following derivations:

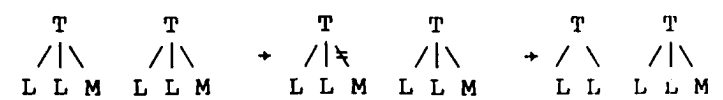
(0.4). a. The fourth tone sandhi (4TS)



b. The third tone sandhi:A (3TS(A))

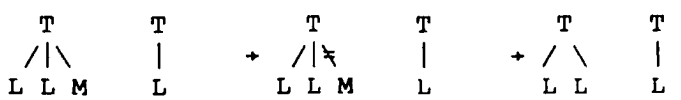


c. The third tone sandhi:B (3TS(B))



d. The second tone sandhi (2TS)

e. The neutral tone sandhi (0TS)



While the implementation rules in (0.3) produce grammatical results in all other cases, they yield outputs, in the cases of 3TS(B) and 0TS, that violate the following OCP related, Mandarin specific WFC:

(0.5). * T T
|\ /|
X X

where X = H, M, or L

Therefore, these two outputs obligatorily undergo the following OCP repairs ((a) for 3TS and (b) for 0TS):

(0.6). a. T T
|\ /|
L L
+
M

b. T T
|\ /|\
L L
+
M

In brief, all the Mandarin tone sandhi processes are fundamentally tone reduction processes, the results of which may be subjected to further modification should they turn out to be violations of certain WFCs. In addition to the rules and derivations presented above, the analysis proposed also contains a theory of tonemes as timing units, a theory of Mandarin syllable weight and quantity, and a theory of the diachronic implications of the analysis.

Examiners:

Dr. Barry F. Carlson, Supervisor (Department of Linguistics)

Dr. Thomas E. Hukari, Departmental Member (Department of Linguistics)

Dr. James Arthurs, Departmental Member (Department of Linguistics)

Dr. Daniel J. Bryant, Outside Member (Department of Pacific and Asian Studies)

Dr. William S-Y. Wang, External Examiner (University of California at Berkeley)

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am grateful to many people here in Victoria who have generously given me time and advice, and shown me kindness, understanding and patience during my doctoral studies.

I extend the deepest gratitude to my committee members, particularly to my supervisor, Dr. Barry Carlson. I enjoyed very much Dr. Carlson's style of supervision. He did not set too many rules for me to follow, but always showed his confidence in me, and was always there when I needed help. Also, I greatly appreciate his very careful perusal of drafts of this dissertation. His extensive and detailed comments on these drafts led to many significant revisions of this dissertation. Dr. Tom Hukari, another member of the committee, also made major contributions to this dissertation. His "relentless" comments on an early draft put me on guard and made me to think twice before making any conclusion. Although I found those comments hard to swallow at the time (but not because I was not convinced), they proved to be extremely insightful and beneficial to my later revisions. The committee member Dr. James Arthurs was also the instructor of my first linguistics course. Thus, I am in debt to him also because it was his qualitative presentation of the basics in this field, punctuated with humorous anecdotes about languages, that led me into the fascinating world of linguistics. As well, my thanks go to Dr. Daniel Bryant, the outside member, who, too, gave valuable comments on the final draft of the dissertation, and to Dr. Richard King, the "ex"-outside member, who had given me excellent advice and guidance before having to leave the committee at the final stage to go on sabbatical.

Very special thanks are due to Dr. William S-Y. Wang, the external examiner. The expertise Dr. Wang provided in the areas of Chinese phonology and historical phonology was of particular importance in the formulation of a number of ideas in the dissertation. Although he was an extremely busy person, he was always available, and although he traveled constantly, he was always reachable. I still remember "bugging" him through electronic mail for materials, advice and information, when he was away on international trips.

I am very grateful to the Department of Linguistics and to its Chair, Dr. Joseph F. Kess, for the financial support both in the form of the fellowship and the teaching/research assistantship, and to Dr. Richard King for hiring me as a Teaching Assistant for Chinese. Without such financial assistance, there would not have been this dissertation.

The following people have also helped me in various ways during my doctoral studies: Professors Tom Hess, Leslie Saxon, John Esling, Barbara Harris, the general office secretaries Darlene and Gretchen, and my friends Bob, Robynne and the late Nancy Swailes. The Swailes welcomed me to Victoria and very kindly helped me through the critical period of cultural shock.

Last but not least, I thank my family: my husband, my son, my mother, my brother and sisters. My son always showed me his understanding when I was busy with the dissertation and could not have time for him, even though he did not really understand why mom had to work on "the book".

I assume full responsibility for any mistakes in this dissertation.

To my father, Hui Lin, in loving memory.

Chapter I

INTRODUCTION

The most well-known feature of the Chinese language is probably that it is a language with phonemic tones. Although Chinese tones (or tones in languages in general) have suffered relative neglect in the history of generative phonology, the last decade has witnessed a steady growth of interest in their study. Largely inspired by the recent theoretical advances made in generative phonology, in particular, by the advent of non-linear phonological theories such as autosegmental phonology (Goldsmith 1979, 1990, Lieber 1987), metrical phonology (Liberman and Prince 1977), and prosodic and sentence phonology (Selkirk 1978, 1984, 1986; Kaisse 1985, Kaisse and Zwicky 1987, McCarthy 1982, McCarthy and Prince 1986, Nespor and Vogel 1983, 1986, Vogel 1984, Zwicky 1985), this increased interest has resulted in a sizable number of studies being conducted with an exclusive focus on Chinese tones.

Within the fields of prosodic and sentence phonology, for instance, Chinese tones and their well-noted sandhi processes have been re-investigated, this time, from a more global perspective. Earlier, Cheng (1973) had made the bewildering observation that although Mandarin tone sandhi processes always occur within an isolated string of two syllables, their application becomes something of a puzzle when more syllables are added to the string. Specifically, when the string becomes longer, the application of the tone sandhi rules becomes somehow controlled by the surface syntactic structure. Despite the intriguing nature of Cheng's observation, the syntax-tone relation has remained elusive and intractable for two dec-

ades, due apparently to the lack of a viable tool for its study. This situation, however, has greatly changed since the theories of prosodic and sentence phonology came into being. Armed with these theories, many studies have been conducted, probing into the syntax-tone relation not only in Mandarin but in other Chinese languages as well (Chen 1990, Cheng 1987, Chung 1992, Hung 1987, Lin 1990 and 1991, Shen 1986, Shen 1988, Shih 1986, Selkirk and Shen 1988 and 1990, Wright 1983, and Zhang 1988). These studies have added much to the current understanding of the interaction in question. For instance, one major understanding that has been advanced by these studies is that Chinese tone sandhi rules do not exactly apply on syntactic brackets, as they had been thought to, but rather, on some kind of an intermediate prosodic structure completely or partially projected from the syntactic structure.

Besides being approached externally in terms of the domains of their sandhi behavior, Chinese tones have also been tackled from within. The age-old question of how Chinese tones should be most optimally represented by a phonological featural system is raised again in light of the newly-developed autosegmental phonology. Not exactly surprisingly, the inquiry in this direction has led to the conclusion that Chinese tones are essentially *autosegmental* (Yip 1980a). A concept newly developed by Goldsmith (1979), the term "autosegmental" refers to the state of a phonological feature not behaving completely in accordance with the rest of the features, but showing independent or autonomous traits. Since Yip's pioneer work in 1980 (i.e. 1980a), there have been a number of studies on the internal structure of Chinese tones (Yip 1989a,b, Bao 1990, Shen 1985, Packard 1989, Zee 1991).

In spite of the increased interest and the large number of studies that have emerged as a result of it, a review of the literature still reveals some problems (or gaps) in the study of Chinese tones. Generally speaking, tones can be approached

from at least two different perspectives. They can be studied in terms of their internal representation, or they can be tackled in terms of their sandhi processes. The sandhi processes can themselves be approached in two fashions. They can be analysed to answer the question of where they actually occur in a phrase or sentence. Alternatively, they can be examined internally to see what exactly happens when they actually occur; namely, what the exact nature of these sandhi processes is, or why tone sandhi processes occur at all within those specifiable domains.

However, for some reason or other, the overwhelming majority of the recent studies on Chinese tones focus exclusively on the problem of the domains of Chinese tone sandhi, conducted within the prosodic and sentence phonology. Relatively few examine the internal representation of Chinese tones. In addition, none of these studies has attempted to explain why Chinese tone-sandhi processes occur. Although it may be true that the studies of the tone-sandhi domains may, to a large extent, ignore the internal structure of the pertinent tone, the exact shape of the tone sandhi rule for which the domains of application are sought crucially depends on an optimal representation of the pertinent tone. Such a relation between representation and processes is due to an understanding recently reached in non-linear generative phonology. According to this understanding, phonological representation is more fundamental than phonological processes in the sense that configuration of adequate rules of phonological processes (such as tone sandhi rules) crucially depend on an optimal representation of the elements (such as tone) involved in the phonological processes.

Another general problem in the studies of Chinese tones is that most of the studies cover in theory all of the tonal systems of the Chinese languages. Although such a coverage within a single work of the varied (though closely relat-

ed) languages may not in itself present a problem, it does point to a lack of serious efforts in more focused, in-depth studies of just one of the languages covered under the proposed theories. One major characteristic of these recent studies is that they are all theory-driven, which is not surprising given that they are all inspired by the relatively recent availability of non-linear phonological theories. Being theory-driven is again not questionable in linguistic studies. However, when a study is both theory-driven and at the same time targeted at a large array of various (though closely-related) languages, the validity of the theory emerging from it may need to be further verified. One natural way to verify such a theory is, of course, to conduct in-depth studies of the individual languages covered by it.

The general purpose of this dissertation is defined accordingly. It is to conduct an in-depth study of one of the Chinese languages: Mandarin. Specifically, the purpose is twofold: to investigate (1) the nature of Mandarin tones, namely, the representational aspect of Mandarin tones, as well as (2) the nature of Mandarin tone sandhi processes.

As implied in the above discussion, the representational aspect of Mandarin tones has not received substantial, close-up study in generative phonology. Furthermore, no studies have dealt with the exact nature of the Mandarin tone sandhi processes, though much has been theorized with regards to where, given a string of syllables, they will occur. However, the fact that Mandarin tone sandhi varies in nature from tone sandhi in other Chinese languages has been noted in several studies (Chen 1990, Zhang 1989 among others). But what exactly is meant when Chen remarks that Mandarin tone sandhi processes are basically phonetic in nature while those in other Chinese languages are not? To answer this question, it is obviously necessary to know first of all the true nature of Mandarin tones and tone sandhi processes.

The organization of this dissertation is as follows: the remaining sections of this chapter provide some preliminary information about the focus of this study and the data-transcription conventions adopted in it. In Chapter 2, a general survey of the literature on tonal representation is conducted. One study briefly reviewed (see § 2.6) has special importance to the present study. It is Goldsmith's (1979) *Autosegmental Phonology*. In fact, the review of Goldsmith's work is also meant to introduce the basic theoretical framework used in the analysis proposed in this dissertation. Within the literature on Chinese tonal representation, Yip's (1980a) work stands out as virtually the only one that is both comprehensive and conducted within the current non-linear phonology. However, it is demonstrated in Chapter 3 that Yip's four-level, register-feature tonal system lacks motivation from Mandarin. Further, in Chapter 4, 5 and 6, it is shown that Mandarin tones are best analysed in terms of a three-level system, with which a unified analysis of the Mandarin tone sandhi rules becomes for the first time possible. The three-level system is spelled out in Chapter 4. Chapter 5 discusses an important theoretical concept, a concept concerning segment timing. This concept constitutes an integral part of the theory proposed in Chapter 4. Chapters 6 probes into the nature of the Mandarin tone sandhi processes, and a unified account of them is proposed in it. Further support for the proposed Mandarin tone sandhi analysis is drawn from elsewhere, including historical evidence that involves other Chinese languages. The discussion of this further evidence is provided in the last chapter.

1.1 The Chinese Languages

There has been much controversy concerning whether to speak of the Chinese language in the singular or the Chinese languages in the plural. It is generally known that Chinese has eight branches. The controversy is hence over whether to

regard the eight branches as dialects of a single language or languages themselves. The major argument for the latter position is based on the Criterion of Mutual Intelligibility (Steinbergs 1987). By this criterion, two linguistic entities are dialects if they are mutually understandable; otherwise, they are languages. Since the eight branches of Chinese are mutually unintelligible, they should then be regarded as languages.

An argument for the opposite position is that these branches of Chinese are mutually unintelligible primarily because they have significantly different phonological structures. When it comes to morphology and syntax, relatively little difference is found among them. Other arguments for the opposite position are drawn on language-external factors such as that all these branches share the same written form, and that all the speakers of Chinese are of the same ethnic origin, sharing the same history and culture (cf. Baxter 1992, DeFrancic 1984, Ramsey 1987).

I do not wish to join this argument here. In this particular study, I will use "the Chinese languages" in the plural form. This choice is made in light of the considerations that the present study is a phonological study (and not a morphological or syntactic one), and the fact that, as just mentioned, these branches of Chinese are significantly different phonologically. The choice of the plural form also reflects a practical consideration: I can then refer to the sub-branches of one of these Chinese languages as dialects, thus avoiding unnecessary confusion that is likely to result if the term "dialect" is used for items at both levels. The following illustrates my use of the terms "dialect" and "language".

(1.1). The Chinese Languages and Dialects

	Languages	Dialects
1.	Mandarin (Northern Dialect)	Mandarin (Standard Dialect)
		Mandarin (Beijing)
		Others
2.	Wu	Shanghai
		Others
3.	Gan	Nanchang
		Others
4.	Xiang	Changsha
		Others
5.	Hakka(Kejia)	Meixian
		Others
6.	Southern Min	Xiamen(Amoy)
		Others
7.	Northern Min	Fuzhou
		Others
8.	Yue	Guangzhou(Cantonese)
		Others

1.2 Mandarin as Standard Chinese

The list of the Chinese languages and dialects given in the previous section reveals one problem. That is, the term *Mandarin* has been used simultaneously to mean three things: Mandarin as one of the Chinese languages, Mandarin as the Standard Dialect, and Mandarin as any individual dialect (such as Beijing) of the Mandarin language. A question thus arises here as to what I mean by claiming that the present study focuses on Mandarin. Although the demarcation may not be terribly necessary when a morphological or a syntactic study is concerned, it is rather important in a phonological study. Especially in terms of tonal categories, tonal values and tone-sandhi phenomena, Mandarin as a language can be quite different from other Chinese languages. For instance, Mandarin has categorically four tones, but Cantonese, a Yue dialect, has as many as nine. Mandarin tone sandhi processes have been observed to be basically phonetic in nature, while tone sandhi processes in other Chinese languages may relate to morphology. Even the dialects within the Mandarin language can differ greatly in their tonal shapes and tone sandhi processes. Although tone sandhi processes in the language of Mandarin may share the common feature of all being phonetic in nature, they nevertheless may be quite different from dialect to dialect.

In this study, the focus is truly on Mandarin as the Standard Dialect (MSD). Taken as the standard Chinese in both the People's Republic of China as well as the Republic of China (i.e. Taiwan), MSD has been officially defined in the People's Republic of China as a dialect whose pronunciation is based on that of the Beijing dialect (though not identical to it), and whose morphology and syntax are based on the Mandarin language as a whole. This standard dialect is the dialect used by the various broadcasting media, and is widely in use everywhere in China. In the non-Mandarin areas, it is normally used as a "second language". In this dis-

sertation, I will use the term Mandarin specifically for MSD, unless otherwise noted.

1.3 The Minimal Domain

As I mentioned earlier, linguists have found that Mandarin tone-sandhi processes do not just occur wherever their structural descriptions are met; rather, they occur in some designated fashion on a prosodic structure built completely or partially on the surface syntactic structure. One area of study on Mandarin tones is thus to determine exactly how this intermediate structure is built on a given string of syllables, so as to find out where exactly on this string these tone sandhi processes actually occur. Earlier, I have also mentioned that one of the specific purposes of this dissertation is to explore the nature of Mandarin tone sandhi. Such a focus of investigation means that I am interested in why Mandarin tone-sandhi processes happen rather than where they happen. In other words, the present investigation is limited to the minimal domain of Mandarin tone sandhi.

What then is the minimal domain of Mandarin tone sandhi? Can one characterize it? And how? These questions are easily answered if the strings under scrutiny contain only (isolated) two-syllable words or phrases (or three-syllable words or phrases in the case of the second-tone sandhi to be discussed in § 6.2.4), since these strings consist of the minimal domains themselves. How then can one define the minimal domain on longer strings? Fortunately, this definition has been provided by studies within prosodic and sentence phonology. In these studies (Shih 1986 many others), this minimal domain for Mandarin tone sandhi has been defined as a prosodic foot.

According to Shih, the prosodic foot is built by the following three rules (p. 110):

(1.2). Foot Formation Rules

a. Immediate Constituency:
Link immediate constituents into dissyllabic feet.

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

c. Triple Meter:
Join any leftover monosyllable to a neighboring binary foot according to the direction of syntactic branching to form a superfoot.

Although controversy still exists in the exact details of the formation of a prosodic foot, this controversy need not concern us here. The present study will be limited to the exclusive domain of a prosodic foot, regardless of how exactly such a foot is defined, but assuming simply that it is already optimally defined in prosodic and sentence phonology.

There is one more clarification that needs to be made here regarding the minimal domain of a foot. As Cheng (1973) has observed, the domain of Mandarin tone sandhi becomes larger as the tempo of speech increases (cf. § 7.1). This means that given enough speed, all the syllables (no matter how many) in a multisyllabic sentence may be compressed into one single domain. In order to keep the present study within its proper limits, I will consider primarily the minimal domain of a foot under slow or normal speech rate.

1.4 Data Transcription

Unless otherwise noted, the Mandarin data presented in this dissertation will be transcribed in the writing system of *Hanyu pinyin* (henceforth *pinyin*), the standard Romanized writing system officially adopted in 1958 in the People's Republic of China. For the last ten or so years, this system has become quite popular among linguists of Mandarin. In fact, it has virtually become the only system of

transcription in Mandarin linguistics. In this dissertation, not only will my own Mandarin data be presented in this system, Mandarin data cited from earlier studies presented in other systems (e.g. Chao, 1968) will also be presented here converted to *pinyin* - unless otherwise noted. The International Phonetic Alphabet symbols will also be used alongside of *pinyin*, where their use becomes necessary. Following tradition, these symbols will, however, be included in slashes // (or square brackets []).

Chapter II

A REVIEW OF PREVIOUS STUDIES ON TONE

2.1 Introduction

One of the most important theoretical advances made in recent generative phonology lies in its shift of focus from rules to representation. The idea is that only if the representation is correct can optimal rules be formalized; the reverse is, however, not true. The literature review conducted in this chapter focuses accordingly on the phonological representation of tones, especially, of Chinese tones.

A number of relevant issues have been explored in the literature on tonal representation. Among them is, first of all, what constitutes a tonal feature. Specifically, the question concerns: should the universal set of tonal features contain unitary *contour* features such as [rising] and [falling], or should it contain only *register* (or level) features such as [high] and [low]? In the current non-linear phonology, this is a question of whether the contour tones should be regarded as indivisible units describable in terms of unitary contour features, or be seen as being composed of a mere sequence of level tonemes describable only in terms of register features. These two positions are illustrated below using the Mandarin rising tone as an example.

(2.1). *Example*

má

Representation

a. LH b. [+rising]

Traditionally, the (b) type of approach was widely adopted (e.g. Pike 1948, Elimelech 1974, Schuh 1978), particularly with regards to Chinese tones (Wang 1967). However, this approach came under serious attack in 1969 by Woo, who proposed the (a) type of approach instead, basing her proposal on evidence from Chinese. The central claim of the (b) treatment was further challenged by Goldsmith in his *Autosegmental Phonology* (1979),[1] with evidence from African tones. Yip (1980a), following Goldsmith, has developed the first non-linear representation of Chinese tones, and just as in Goldsmith, her representation treats the contour tones as being composed of level tonemes.

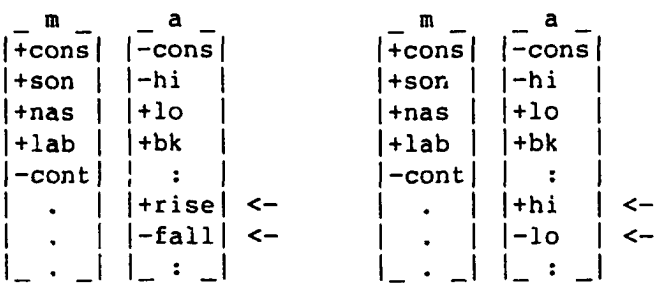
In recent years, however, a third approach has been proposed by Yip (1989a,b) in an attempt to solve the problematic cases that cannot be reasonably handled by a level-primitive-only analysis. The third approach is, in a sense, a compromise: it is less strong than either of the two former positions, with the basic claim that although at one level, contour tones are best accounted for in terms of sequences of level tones, they may be describable at another level only if treated as primitive units.

Besides the issue of the contour tone treatment, another important focus of studies in the area of tonal representation concerns the question of how tonal features can best be configured in the phonological representation. This is the question which eventually led to Goldsmith's (1979) revolutionary modification of the generative approach to phonological feature representation. The major question addressed in Goldsmith is whether the tonal features in some languages (or rather certain features, not necessarily tonal, in some languages) should be granted autonomous status and represented separately from the rest of the features. The raising of this question is in itself quite unconventional, as it implies that features after all may not necessarily always be bundled together within the same sequence of feature matrices.

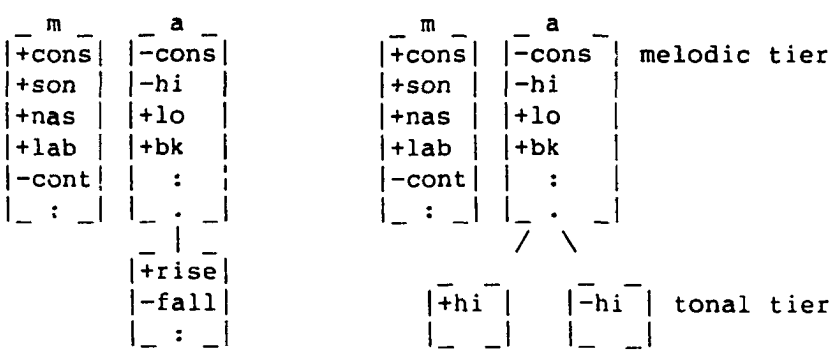
This question, however, has been considered resolved after Goldsmith's insightful work and many subsequent studies within the framework developed by him. These studies have argued cogently and repeatedly with evidence from various languages that tonal features (as well as some other non-tonal features) often need to be represented separately from the rest of the phonological features. Although the debate over the possible autonomous status of the tonal features in the representation is largely settled, the possibilities brought about by such a position lead to new questions to be answered. One such question is: if a tone is autosegmental in a language, what then is the tone bearing unit (TBU) in that language? Is it, say, the vowel or the syllable? Again, let me take the Mandarin second-toned morpheme *má* "hemp" for an example to illustrate the difference (The arrows of the shape <- are used simply to highlight the features relevant in the present discussion.):

(2.2). Contour Treatment (i). Level Treatment (ii).

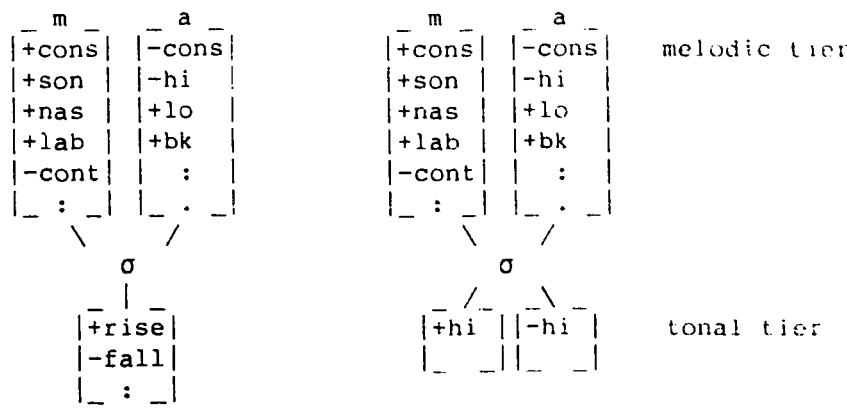
a. segmental treatment



b. autosegmental treatment: vowel as TBU



c. autosegmental treatment: syllable as TBU



As illustrated above, there are in theory at least four possible treatments within the non-linear approach (b and c). If one rejects the treatment of contour tones as consisting of level tones only, there still remain two possibilities (i.e. bii and cii).

Other issues concerning tonal representation include how many levels should be distinguished in the tonal system of a given language or languages. It seems that in most of the African and North America tonal languages, two is the answer. For Chinese-like oriental languages, however, several theories have been proposed, covering a wide range from three levels in Halle and Stevens (1971), four levels in Yip (1980a) and Gruber (1964), to five levels in Sampson (1969), Maddieson (1972),[2] Wang (1967), and Woo (1972).[3] For a given individual language, the hypothesis can be quite diverse as well. For instance, Mandarin tones have been analyzed as contrasting three levels (Woo 1972, Zee 1991), four levels (Yip 1980a), and five levels (Chao 1968).

In the rest of this chapter, I will briefly review the major studies on tones, particularly with respect to their positions in the issues mentioned above. It should be noted, however, that all studies do not bear on all the issues, which means that, in the course of the review, I will focus on one or another of the issues where it becomes pertinent for discussion.

2.2 Pike (1948)

One major characteristic of Pike's study of tones that is of special concern here is that Pike is probably the earliest proponent of the treatment of contour tones as units. In his now classic work on tone, Pike clearly distinguishes two types of tonal languages, level-tone languages such as Mixteco (a language of southern Mexico), and contour-tone languages such as Chinese. Pike argues that while the contour tones in level-tone languages may be sequences of level tonemes, in the description of the contour tones of the contour-tone languages, "the basic tonemic unit is gliding instead of level (p. 8)." Pike gives further comparison between a contour-tone language and a register-tone language, and observes that in a contour-tone language (Pike 1948, p.8),

1. The unitary contour glides cannot be interrupted by morpheme boundaries as can the nonphonemic compounded types of a register system.

2. The beginning and ending points of the glides of a contour system cannot be equated with level tonemes in the same system, whereas all glides of a register system are to be interpreted phonemically in terms of their end points.

3. In the printed material examined, contour systems had only one toneme per syllable, whereas some of the register-tone languages...may have two or more tonemes per syllable.

Based on these observations, Pike concludes that

> "in a pure contour system, the glides are phonologically unitary, morphologically simple, and not structurally related to a system of level tonemes; the glides are minimum structural units of length in words and syllables." (p.8)

One feature in Pike that is worthy of remark here is his use of the term *toneme*, which is used to refer to a minimal component of a syllable tone. His unitary treatment of contour tones is also reflected in his use of this term: unlike

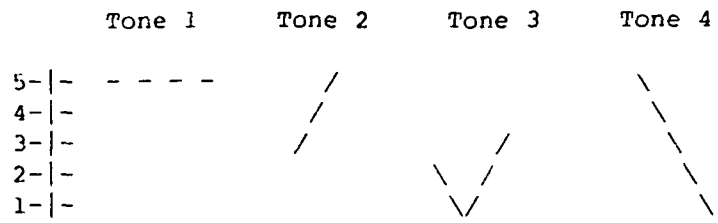
the minimal tonal component defined to be always level in current generative phonology, his toneme may be an atomic contour in itself. Thus, according to him, a tone of a syllable typically contains one level toneme in a register-tone language, but a single contour toneme in a contour-tone language. However, just as a tone in a register-tone language may occasionally contain two level tonemes, resulting in a simple glide, so may a contour tone in a contour-tone language contain two contour tonemes, resulting in a complex contour tone (i.e. a concave or convex tone).

For some reason or other, Pike's term *toneme* as a minimal component of a tone has not been adopted among later tonologists, although it clearly allows for a useful (and necessary) distinction to be made more explicit between the tone at the syllable level and that at the sub-syllable level. Normally, tonologists simply use *tone* for units of both levels. The reluctance to use the term "toneme" is probably due to the great complexity in the possible internal structure of what may be called a tone in the various tonal systems in human languages.

2.3 Chao (1930, 1968)

It is probably safe to say that Yuen Ren Chao is the forefather of contemporary linguistic studies of Chinese. One particularly outstanding contribution he has made to these studies is the scale of five pitch levels he developed for tonal representation. An interesting phenomenon one will not fail to notice in reading literature on Chinese tones is that phonologists of Chinese may each argue for a quite different theory of Chinese tonal analysis, but when it comes to the introduction of tonal data at the outset of their study, they almost invariably choose Chao's five-level scheme for the task, no matter how much more advanced and insightful their theories may be. In what follows, I will illustrate Chao's scheme using the Mandarin tones as an example.

(2.3).



Chao's scheme, as is illustrated above, divides the human vocal range into five levels, with the first and the fifth marking respectively the lowest and the highest pitches a speaker reaches in normal speech production. Here the term "normal" is critical since the scheme does not deal with unusual pitch levels that, for instance, are raised higher than the fifth in singing or due to affective factors such as anger or joy.

On Chao's scale, the first Mandarin tone (being the first or any other number merely by convention) starts at the top of the speaker's pitch level (marked as level 5) and remains so for a certain period of time. The second starts in the middle (level 3) of the vocal range and rises sharply to the top. The third is a more complex one. It first drops from level 2 to the lowest pitch level (level 1) and then rises again to the middle point. Based on the readings on this scale, Chao refers to the four Mandarin tones by the number composites such as (55) for the first tone, (35) for the second, (213) for the third and (51) for the fourth.

Chao's scale readings also have had impact on the kind of diacritics that have been adopted to mark Mandarin tones. For instance, the four diacritics used in the Romanized Chinese writing system *pinyin* are not totally arbitrary symbols. A comparison between these diacritics (shown below in column (a)) and Chao's scheme reveals that each diacritic has the same contour shape as the shape of the tone represented on Chao's scale:

(2.4).

(a) pinyin	(b) number composites	(c) gloss
mā	(55)	"mother"
má	(35)	"hemp"
mǎ	(213)	"horse"
mà	(51)	"scold"

One unique characteristic of Chao's system is that it is a scalar rather than a binary-feature system. This characteristic distinguishes Chao's from all later analyses of Chinese tones. Another characteristic of Chao's that is worth mentioning is that it is basically an analysis that recognizes contour tones as units. Of course, when Chao's theory was first developed (1930), there was not yet an awareness of a distinction between the two approaches to the representation of contour tones. In fact, no discussion of the issue was made in his work 36 years later (1968). Nevertheless, Chao's theoretical orientation with regard to this issue can be described with fair amount of certainty. As Bao (1990) has remarked,

> "in Chao's numerical representation, the tones are conceived to be single, atomic entities. 53 does not imply that the high falling tone is composed of the height point 5 followed by the mid point 3. 53 is a unitary high falling tone,...and 31, a unitary low falling tone." (p.21)

2.4 Wang (1967)

If Chao's theory is "pre-theoretical", Wang is the first to develop a tonal representation for the Chinese tones, couched in the then relatively new model of generative linguistics. In fact, Wang's study represents the next major work on Chinese tones after Chao, and it pioneers generative research on the then linguistically little known phenomenon called tone.

A departure Wang makes from Chao is that he is an explicit proponent of a unitary analysis of contour tones. Within the generative school, Wang devises the first set of binary features specifically for tonal representation. This set of features contains unitary contour features as well as register features. In all, Wang develops seven binary tonal features, four contour features: [rising], [falling], [contour], and [convex], which are to represent the contours of various tones, and three register features: [high], [central] and [mid], which are to define five levels of tones. The five levels of tones are defined as follows:

(2.5).

	[high]	[central]	[mid]
(55)	+	−	−
(44)	+	+	−
(33)	−	+	+
(22)	−	+	−
(11)	−	−	−

In Wang's feature system, the four tones of Mandarin can be represented as in (2.6). It should be borne in mind in reading (2.6) that Wang's system is intended not just for Mandarin, and therefore, some of the features in (2.6) may not appear as motivated as they actually are:

(2.6).

Tone 1 (55)	[−rising, −falling, −convex, −contour, +high, −central, −mid]
Tone 2 (35)	[+rising, −falling, −convex, +contour, +high, −central, −mid]
Tone 3 (213)	[+rising, +falling, −convex, +contour, −high, −central, −mid]
Tone 4 (51)	[+falling, −rising, −convex, +contour, +high, −central, −mid]

While the motivation behind the use of the features [rising], [falling], [contour] and [high] appears quite straight-forward, his postulation of the rest of the features needs a little explanation. It seems that in order to represent a complex contour tone (i.e. concave, convex, etc.), which contains both rising and falling components, Wang uses the feature [convex] to ensure the desired order of the ris-

ing and falling components. Wang's use of the features [central] and [mid] seems at first sight redundant. Their contrast, however, is to distinguish the tones 22, 33 and 44. A tone with a value of 22, for instance, is [+central, -mid, +high]; one with a value of 33 is [+central, +mid, -high]; and one with a value of 44 is [+central, -mid, -high].

It seems that Wang was not only the first, but also the last to have developed and used contour features explicitly and in a rather systematic way for tonal analysis. Although not long after his work, his approach of treating contour tones as units gave way to a new approach which considers contour tones as special forms of level tones, nonetheless, the plausibility of the former approach has yet to be judged. Recent studies by Yip (1989a,b), for instance, have found that at least at some level in some Chinese languages, contour tones should be treated as units. Similar findings are also made in the present study (cf. § 4.2.4.).

2.5 *Woo (1972)*

The studies I have reviewed so far all share more or less the common feature of representing contour tones in a unitary way. This assumption is first challenged by Woo (1972). Woo's feature system for tonal analysis assumes a total of three features, [high], [modify] and [low], all being register features. These features combine to yield five level tones, just as Wang's level features do. These five levels are shown below (Woo, p.71):

(2.7).

	[high]	[modify]	[low]
(55)	+	-	-
(44)	+	+	-
(33)	-	-	-
(24)	-	+	+
(11)	-	-	+

Although Woo's features are defined in a way quite similar to the register features in Wang, her system differs drastically from Wang's in that it dispenses totally with contour features such as [rising] and [falling]. According to Woo,

> "a system using the features [rise] and [fall] is neither mechanically nor theoretically adequate to describe the tones of all languages. We thus abandon this approach and turn instead to a system which describes contour tones as sequences of pitch heights." (p. 64)

Woo goes against the tradition of her time with the claim that in deep structure, contour tones are made up of level components. Mandarin tones, in particular, become represented for the first time in the same way as some of the African contour tones are represented, that is, in terms of a sequence of level tonemes. In particular, these Mandarin tones contain two tonemes each (Woo, p.75):

(2.8).

	Tone1 (55)	Tone2 (35)	Tone3 (213)	Tone4 (51)
[high]	+ +	- +	- -	+ -
[low]	- -	- -	+ +	- +
[modify]	- -	- -	- -	- -

As shown above, the feature [modify], with its negative value found for all the tonemes, is not distinctive for Mandarin. This means that Mandarin tones can be specified using just the other two features. Thus, the above can also be stated in the following fashion (Woo, pp. 73-74),

(2.9).

Tone1	Tone2	Tone3	Tone4
[+hi][+hi]	[-hi] or [-lo]	[+hi]	[+lo] [+lo]	[+hi] [+lo]

or using the popular abbreviated symbols H, M and L:

(2.10).

Tone1	Tone2	Tone3	Tone4
HH	MH	LL	HL

Many studies after Woo have demonstrated that the representation of contour tones as being composed of level tonemes is absolutely necessary in phonological

description, confirming that Woo's tonemic treatment of contour tones represents an important advance in the phonological study of tones. However, Woo's theory also has problems, the most serious of which results from the theoretical framework her analysis is couched in: it seems her level-toneme treatment of contour tones is in conflict with the theoretical framework. Within the framework of standard generative phonology, Woo finds it impossible to represent in the same feature matrix the sequence of two level tonemes she proposes for the Chinese tones. Let me illustrate the problem with the following two sets of feature matrices:

(2.11).

(a).

m	a
+cons	-cons
+son	-hi
+nas	+lo
+lab	+bk
-cont	:
.	+rise <-
.	:

(b).

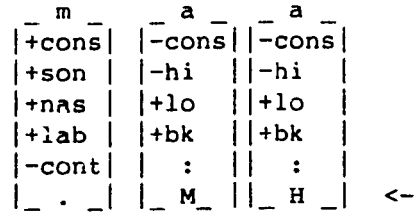
m	a
+cons	-cons
+son	-hi
+nas	+lo
+lab	+bk
-cont	:
.	M H <-
.	

Illustrated in the above are representations of the Mandarin second (35) tone. In (a), a unitary contour approach is used. Under this approach, the rising tone is represented by placing unitary features such as [rising] ([rise] for short) in the feature matrix of the vowel. In contrast to (a), (b) represents the same rising tone by using level tonemes such as L and H. Both (a) and (b) are couched in the traditional segmental approach, which treats tonal features as properties of the vowel (or the tone-bearing sonorant).

While the use of unitary contour features in (a) fits into the traditional segmental approach, a problem arises in the hypothetical (b) in which two feature matrices (abbreviated as M and H) in a designated order are found. However, an important tenet of the standard model she follows is that features in the same matrix are un-ordered. This means that she simply cannot ensure that the M features temporally precede the H features.

Woo's solution to the ordering problem is to "enlarge" the tone-bearing segment, in particular, to assume that there are no single-sonorant syllables in the language. All of the open syllables traditionally regarded as being composed of one single vowel contain two vowels in Woo, so that *ma* becomes *maa*, *di* becomes *dii*, and so forth. With the addition of one vowel, the sequence of level tonemes (two maximum in Woo for underlying representation of Chinese tones) can then be evenly distributed among them, and the case in (2.11b) can then be represented as follows, where the awkward situation seen there disappears:

(2.12).



In such an analysis, Woo apparently assumes the assumption of standard generative phonology that tonal features are but features of the vowels (or the tone bearing sonorants). In other words, tonemes necessarily bear a strictly one-to-one relationship to vowels. This assumption, however, has been quite convincingly rejected by Goldsmith's (1979) autosegmental phonology and Yip's (1980a) autosegmental account of Chinese tonal representation. In fact, it is due to this assumption of the standard generative phonology, rather than anything else, that Woo's number of tonemes has to be limited to no more than two for any underlying Chinese tone. The significance of this two-toneme limitation lies in that no convex or concave tones are recognized in deep structure. Let us take the third Mandarin tone (213) as an example. This tone has three forms, (213) in final position or in isolation, (21) when followed by any other tone than itself, and (35) when followed by itself. Woo's analysis for this tone is different from earlier studies not only because she represents it with level tonemes but also because she assumes the

sandhi form (21) rather than the citation form (213) as the underlying. As has been shown earlier, both Chao and Wang treat this tone as a complex contour (i.e. concave) tone in the underlying representation. The following compares Woo, Chao and Wang:

(2.13). Tone 3 (UR)

Chao	213
Wang	[+rising, +falling, -convex, +contour, -high, -central, -mid]
Woo	LL

Woo's most important argument for her analysis of the third Mandarin tone is that if this tone is posited as LLM in the underlying representation, that is, if its citation form (213) is taken as underlying, "the phonology would have to contain rules which would delete some sonorant segment in these syllables, as well as adjust the pitch specification of the shortened syllables (p. 43). " She gives the following example to illustrate her point[4]:

(2.14).

a. (i). / d a a ng /
L L M

(ii). / d a ng /
L L

b. (i). / d a a ng / →
L L M
↓
ø

(ii.) * / d a ng /
L M

In Woo, an utterance-final (or citation) syllable in the third tone contains necessarily three sonorants so as to bear its three tonemes. Thus, the third-toned word *dǎng* "to obstruct" (represented here in *pinyin*) has the segmental shape of /daang/ (2.14a-i) with three tone-bearing sonorants. When the same third-toned syllable appears before another tone (i.e. not in final position), it has the segmen-

tal shape as in (2.14a-ii) with two sonorants. To derive (a-ii) from (a-i), as Woo points out, one segment in (a-i) has to be deleted together with its toneme. This deleted segment, she further notes, cannot be the last segment /ng/ in view of the surface form of (a-ii); it has to be either one of the vowels. Yet, the deletion of a vowel and its toneme (shown in b-i) yields an ungrammatical tonal shape LM rather than the expected LL (b-ii). Woo argues that the problem will not arise if (a-ii) is taken as the underlying form from which (a-i) is derived via the following insertion rules (p. 45):

(2.15). a. [+son] → M / L L ___

b. $\emptyset \rightarrow \begin{matrix} V \\ \alpha Fi \\ L \end{matrix} / \begin{matrix} V \\ \alpha Fi \\ L \end{matrix} \_\_\_ \begin{matrix} [+son] \\ L \end{matrix}$

where Fi is the set of segmental features.

Therefore, Woo concludes, LL (or 2.14a-ii) rather than LLM (or 2.14a-i) should be taken as the underlying.

However, the problem as pointed out by Woo with the derivation of (2.14a-ii) from (2.14a-i) is but a problem with her erroneous assumption of a strict one-to-one correspondence between the sonorants and the tonemes, rather than a problem with the direction of the derivation (that is, with the derivation of LL from LLM). In fact, the problem would not have emerged if the tonal features had been treated autosegmentally (as they should be) by being represented separately from the segmental features. In an autosegmental representation, the deletion of the tonal features making up the toneme L do not have to be accompanied by the deletion of the segmental features (making up the segment /a/) which are associated with the toneme, and the reverse is also true.

In fact, even if the one-to-one correspondence were accepted, Woo's analysis for the third tone would still be problematic. We have just seen Woo's transcrip-

tion (or representation) of the citation syllable *dǎng* as /daang/, in which there are three sonorants to bear the three tonemes LLM. However, Woo's representation of a third-toned syllable in isolation does not consistently contain three sonorants, as shown below:

(2.16). a. /daang/ "obstruct"
b. /zoou/ "walk"
c. /nii/ "you"

When the syllable is a closed one or contains a diphthong, no problem is posed for Woo's sonorant-toneme correspondence. In each case, she just has to assume that the main vowel in the syllable is a long vowel, which should be represented as VV. Thus, the third-toned syllable *dǎng* in isolation is */daang/*, and the third-toned syllable *zǒu* "walk" in isolation is */zoou/*. A problem arises, however, with the third-toned open syllable *nǐ* "you" (2.16b-i) represented in Woo as */nii/*. As mentioned earlier, Woo assumes what used to be considered a single-vowel syllable to be one with two vowels. This way, she can maintain that contour tones are clusters of level tonemes. However, since it would seem absurd to claim that a syllable such as *nǐ* (in *pinyin*) has three vowels in it (that is, with the shape of */niii/*), she has to leave it to be represented as */nii/*. Yet, with only two vowels in this third-toned citation syllable, its three tonemes involved cannot be properly distributed the way they should in the theoretical framework she uses.

2.6 Goldsmith (1979)

The last few years of the 1970s witnessed a very important theoretical advance in generative phonology. Lying at the centre of the advance is the autosegmental phonology developed by Goldsmith (1979). Autosegmental phonology, though developed out of the standard generative phonology, is fundamentally different from it in major ways.

In the standard phonology laid out in the *Sound Pattern of English*, (Chomsky and Halle 1968, henceforth SPE), there are two major aspects that are dealt with: the rule aspect and the representation aspect. Regarding the representation aspect, the SPE model projects, among others, the following three basic and mutually dependent assumptions:

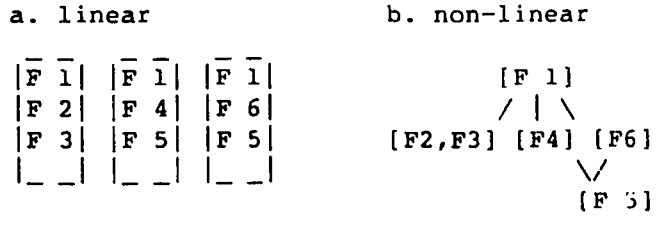
(2.17).

1. *Single Sequence Hypothesis:* This hypothesis says that a human speech continuum consists of one single sequence of segments called consonants and vowels. All phonological rules and processes occur at this single level of segments.
2. *Atomic Hypothesis:* Each of the segments in the sequence is made up of a matrix of feature-value pairs and is atomic in nature; that is, the feature-value pairs are un-ordered in the matrix with regard to one another such that there is no internal structure within each feature matrix.
3. *Linear-Order Hypothesis:* The segments or the matrices of features are ordered in a linear manner.

Eight years after its publication, SPE came to be seriously challenged in Goldsmith's autosegmental phonology. In his 1976 dissertation, published in 1979, Goldsmith argues that SPE fails to adequately accommodate suprasegmental features. His fundamental claim is that a human speech utterance consists of gestures such as tongue movement, lip movement, laryngeal movement, and vocal-cord vibration. These gestures (or articulatory parameters) are in principle autonomous and independent of one another. While they are coordinated so as to produce a synchronized utterance, they do not at all have to start and finish at the same time during the production.

Given this now well accepted argument, the SPE's Single Sequence Hypothesis runs into serious problems. With only one sequence of clear-"sliced" segments, the fact that the gestures do not necessarily have the same duration in speech production cannot be properly configured. To solve this problem, Goldsmith offers his multi-tiered theory, which holds that the phonological representation of a human utterance consists of two or more parallel but autonomous sequences of segments, or *tiers* of segments. In addition, these segments are no longer seen as being arranged in a linear order but, rather, scattered in some designated fashion all over the multiple tiers. The following illustrates the difference between the two approaches:

(2.18). Representation of features



The non-linear, multiple-tier configuration of the segments shown in (b) means clearly the rejection of the SPE's Atomic Hypothesis as well, since now features are regarded as generally autonomous (they are therefore autosegments), and each may potentially stand out from the rest of the features to form a tier of its own. Or, looked at from another point of view, any of the autonomous features may theoretically remain autonomous in the phonological system of a particular language. This argument is clearly expressed in Goldsmith (1990):

> "One of the fundamental aspects of autosegmental representation is the autonomy granted to the segments on each tier--the etymological source of the term 'autosegmental', in fact." (pp. 27-28)

One crucial mechanism in autosegmental phonology is the linking of the multiple tiers by what Goldsmith calls association lines. The idea of the association lines is based on the assumption that in the unmarked case, the autosegments in the tiers are not paired off or linked in the underlying representation; rather, they are associated to one another by certain well-formedness rules and principles. Regarding the representation of tone, Goldsmith (1979), for instance, proposes the following Well-formedness Conventions (WFC):

(2.19). Well-Formedness Conventions

(1) All tone-bearing units are associated with at least one tone.

(2) All tones are associated with at least one tone-bearing unit.

(3) Association lines do not cross.

With the projection of the features into autonomous tiers and with the conceptualization of the association lines, it becomes now possible for features not to be mapped in a one-to-one fashion. This possibility means an immediate solution to the tonal feature placement problem seen in Woo's analysis. If one adopts Woo's assumption that a contour tone consists of a series of level tonemes, the abstraction of tonal features away from segmental features makes it possible to maintain the desired order of the tonemes, and to allow a single vowel to bear these ordered tonemes simultaneously.

In fact, like Woo, Goldsmith also argues for the position that contour tones should be treated as a sequence of level tonemes. Some very convincing evidence is cited in Goldsmith to support this contention. However, different from Woo, Goldsmith's evidence is drawn from African languages without any mention of the Chinese languages (or of any of the oriental tonal languages). One important piece of evidence he gives[5] for the rejection of contour features is provided below.

In an African language called Igbo (Green and Igwe, 1963; cited in Goldsmith 1979, p.24), there is a kind of sentence form called "I Main". In such sentences, the verb stem is always low-toned in the unmarked case, but the pre-verbal nominal subject always changes its final tone, which is underlyingly high and level, into a contour falling tone. Following convention, the diacritics ˊ, ˋ, ˆ, and ˇ are used exclusively in the description of the Igbo data to stand for high, low, falling, and rising tones respectively:

(2.20). *In citation*

ékwé "Ekwe (a name)"
àdhá "Adha (a name)"
cì "carry"

In "I Main" sentences

Ékwê cì àkhwá.
"Ekwe was carrying eggs."

Àdhâ cì àkhwá.
"Adha was carrying eggs."

The change from a high level tone into a contour falling one on the final syllable of the subject is obviously caused by an anticipatory assimilation of the tonal feature of the following verb. The crucial question raised by this observation is that of how the resultant falling tone is to be represented phonologically. One response is to use a unitary feature [+falling] for it. Goldsmith argued that such a treatment would fail to capture the observation that the second half of the tone resembles, if not identical to, the tone on the following verb. This observation is, on the other hand, easily accounted for if the contour tone is represented as a sequence of two tones, one high and one low, with the low one being mapped to the vowel in question from the low toneme of the following verb by an association line.

Success with Goldsmith's autosegmental phonology has led to tremendous developments in generative phonology. The current non-linear phonological theory and practice have their roots in autosegmental phonology. As mentioned previously, one major impact autosegmental phonology has had on generative phonology manifests itself in terms of a major shift of its general focus of inquiry from a preoccupation with phonological rules to an emphasis on phonological representation. It is now a generally-held assumption that valid characterization of phonological processes in terms of phonological rules crucially depends on an adequate representation of the phonological units and not, as used to be believed, vice versa.

The success of Goldsmith's autosegmental phonology, especially its fairly optimal treatment of suprasegmentals such as tone, has led to a revival of interest in analyzing Chinese tonal phenomena. One particularly noted study spurred by such an interest is Yip's 1980 MIT dissertation entitled *the Tonal Phonology of Chinese (1980a).*

2.7 Yip (1980a)

Yip (1980a) is the first non-linear account of Chinese tones. Her study is significant not only because it is the first non-traditional analysis of the Chinese tones, but also because it has had a strong impact on other non-linear phonological studies of Chinese tones. Her framework of analysis has been adopted partially or completely in a number of studies, and further theories have been developed in these studies based on her framework (e.g Packard 1989, Pulleyblank 1986, Shih 1986, among others).

Yip's study has made an important contribution to our understanding of Chinese tones with her convincingly-presented argument that tones in Chinese should

be treated autosegmentally. In fact, Yip's treatment in general conforms to the native speaker's intuition implied in many previous studies that tones in Chinese behave as if they have a life of their own. Such intuition, for instance, is seen in Wang's (1967) essentially segmental analysis of tones:

> "[Chinese] segmental features are usually not relevant in the various types of tone sandhi; that is to say, the interaction of tones in a sequence is independent of the nature of the segments which occur with the tones... It is preferable to formalize the tone features differently from the segmental features and regard them as features of individual syllables." (p.95)

Although Wang does not spell out just how the tonal features should be configured as features of the individual syllables, his remark nonetheless points to the "deviant" behavior of Chinese tones.

Not only does Yip differ from her predecessors with her autosegmental rather than segmental account of Chinese tones, but she is also unique with her system of tonal features. Essentially, Yip's system consists of two (rather than three as seen in Woo and Wang) level tonal features, called [upper] and [high], with the latter being renamed to [raised] later in Pulleyblank (1986). Yip names the former *Register Feature* and the latter *Tonal Feature*. Within the autosegmental framework, each of the two features claims a separate independent autosegmental tier, rather than being placed in the same feature matrix on the same tier. In addition, these two features overlap in their coverage of the pitch range, rather than each claiming a portion of the pitch range. The Register Feature, for example divides the whole pitch range into two even portions, and the Tonal Feature, divides each of the two portions into two sub-portions. These two features interact to yield a

total of four pitch levels, as shown below: (Chao's 5-level scale is also given here as a reference.)

(2.21).

	[+high]	5
[+upper]		
	[-high]	4
		3
	[+high]	2
[-upper]		
	[-high]	1

Or, the four level tones may be represented as follows:

(2.22). [+upper, +high] or [+upper, H]
[+upper, -high] or [+upper, L]
[-upper, +high] or [-upper, H]
[-upper, -high] or [-upper, L]

Within this framework, Yip gives the following representations for the four Mandarin tones:

(2.23).

Tone 1	Tone 2	Tone 3	Tone 4
[+upper]	[+upper]	[-upper]	[+upper]
σ	σ	σ	σ
H H	L H	L L	H L

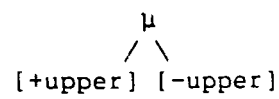
As shown clearly above, Yip follows the autosegmental school by treating contour tones as being sequences of level tonemes, and rejects primitive contour features such as [fall] and [rise].

One special characteristic of Yip's analysis lies in her constraint on the generative power of her analysis. According to Yip, all previous analyses of Chinese tones had been too powerful, over-generating tones that never occur in actual languages. To avoid this, Yip put a condition (given below) on her Register Feature which requires the Register Feature ([upper]) to remain constant in value within the same morpheme:[6]

(2.24). The Register Feature remains constant over the morpheme.

By virtue of this constraint, a tone in the following shape would be eliminated from the grammar:

(2.25).



where μ = morpheme

With this condition on the Register Feature, Yip's analysis generates four and only four contour tones, which are given below:

(2.26).

	rising	falling
[+u]	[-hi][+hi]	[+hi][-hi]
[-u]	[-hi][+hi]	[+hi][-hi]

where u = upper
hi = high

The uniqueness of Yip can be shown further in a comparison of her work with previous analyses. It can be done, for instance, if Yip and Woo are compared. The major differences between the two are obviously due to the fundamental differences between the segmental versus autosegmental theories they each follow. For instance, Yip's tonemes, no matter how many, are linked to the same tone-bearing unit (TBU), which can be a single vowel, whereas Woo's bear a one-to-one relationship to their tone-bearing units.

Another difference between Yip and Woo is that Yip postulates two overlapping features while Woo postulates three more or less discrete ones:

(2.27).

Woo		Yip	
	[high]		[upper]
	[modify]		[high]
	[low]		

With two overlapping features, Yip defines four level tones, while Woo defines five with her three discrete features. In terms of the treatment of Mandarin tones, however, Yip assumes a four-level contrast, while Woo assumes three:

(2.28).

Woo			Yip	
	H	([+high, -low])		[+upper, H]
	M	([-high, -low])		[+upper, L]
	L	([-high, +low])		[-upper, H]
				[-upper, L]

In spite of the fundamental difference between Yip and Woo, they resemble each other in a few important ways. Besides the fact that they both treat contour tones as being composed of level primitives, their most conspicuous similarity lies in the shape of the tonemes for Mandarin tones used in both studies. Just as Woo does, Yip uses pairs of tonemes for Mandarin tones, though Yip's pairs are associated with a single TBU while Woo's are associated with two. The following shows this identity using Mandarin syllable *ma* as an example:

(2.29).

	Tone 1 (55)	Tone 2 (35)	Tone 3 (213)	Tone 4 (51)
	[+upper]	[+upper]	[-upper]	[+upper]
	|	|	|	|
	ma	ma	ma	ma
	/\	/\	/\	/\
Yip:	HH	LH	LL	HL
	maa	maa	maa	maa
	||	||	||	||
Woo:	HH	MH	LL	HL

Another similarity between Woo and Yip that is worth mentioning here lies in the fact that both handle the third Mandarin tone (213 in Chao's theory) in a non-traditional way, that is, as underlyingly a low level tone rather than a complex contour tone.

Although Yip's analysis is essentially correct in terms of its recognition of the need to treat Chinese tones autosegmentally, there are nonetheless problems with it. However, I will not go into these problems here, as a later chapter (i.e. Chapter 3) is devoted totally to their discussion.

2.8 Yip (1989a,b)

Before I address Yip (1989a, b), let us first have a look at an interesting phenomenon with regards to the use of contour versus level tonal features in the study of the tonal systems. It has been observed that Asia and Africa are the regions where tonal languages are mostly found. However, it has also been found that the tonal systems of the Asian languages are typologically quite distinct from that of the African languages. Due to this distinction,

> "the [contour-feature approach] is the one used most frequently by people working with Oriental languages. The [level-feature-only approach] is the one used by most of the people working with African languages ..." (Woo 1972, p.24)

Today, this difference in approach remains, though in a more complex manner. On the one hand, linguists working on tonal systems of the African languages tend to apply implicitly or explicitly the level-feature-only approach, and there is more of a general consensus among these linguists on the choice of this approach. Their data from the African languages are generally analyzed fairly neatly and successfully with the level-feature-only approach.

However, a comparable situation is not found in the studies of the tonal systems of languages such as Chinese. In these studies, linguists are divided; there are followers of both camps. Those who adopt the level-feature-only framework (e.g. Woo 1972 and Yip 1980a) argue that Chinese-like tonal systems are not fundamentally different from those found in African languages, and therefore, should be treated likewise with the level-feature-only approach. As mentioned earlier, this theoretical position has enjoyed good acceptance among phonologists of Chinese. In spite of this, there is one thing that cannot be denied: the same degree of neatness and success achieved by the use of the level-feature-only approach in the

studies of African tonal systems cannot be found in similar studies of Chinese-like tonal systems.

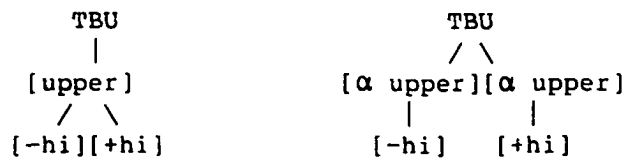
This is probably the reason why earlier phonologists working on Chinese tonal systems are reluctant to accept this level-feature-only approach. Also, this is probably why, after being an advocate of this approach for almost a decade (since 1980), Yip in 1989 turned back to the more traditional perspective for a possible solution to the problematic cases in Chinese. She then proposed what one may call a third approach, one that recognizes the dual possibilities of contour tones behaving as atomic units as well as acting as decomposed sequences of level tonemes. Now, let us turn to Yip's analysis proper.

Yip's (1989a,b) studies basically argue for the position that contour tones in Chinese may associate as melodic units.[7] Using techniques from the theory of feature geometry, Yip distinguishes two kinds of contour tones, branching tones and tone clusters:

(2.30). a. Branching tone b. Tone cluster

According to Yip, contour tones of type (a) are typical in Chinese-like tonal systems, but the cluster type (b) occur primarily in African languages. Notice that the two trees above are in abbreviated forms. If spelled out in feature matrices, the trees look like the following:

(2.31). a. Branching tone b. Tone cluster



What is innovative about the branching type lies apparently in the addition of a unitary contour level between the tonemes and the TBUs. This unitary level makes it possible to represent the contour tone as a unit as well as a sequence of level tonemes (although, as Yip clearly notes, this treatment does not exactly entail the use of contour features).

Yip's major task in her two studies is to argue for the branching type whose existence had been denied since Woo. As may be predicted, Yip's argument relies primarily on evidence from dialects of Chinese. Here, let us have a look at some of her evidence for her argument.

2.8.1 Evidence from Danyang

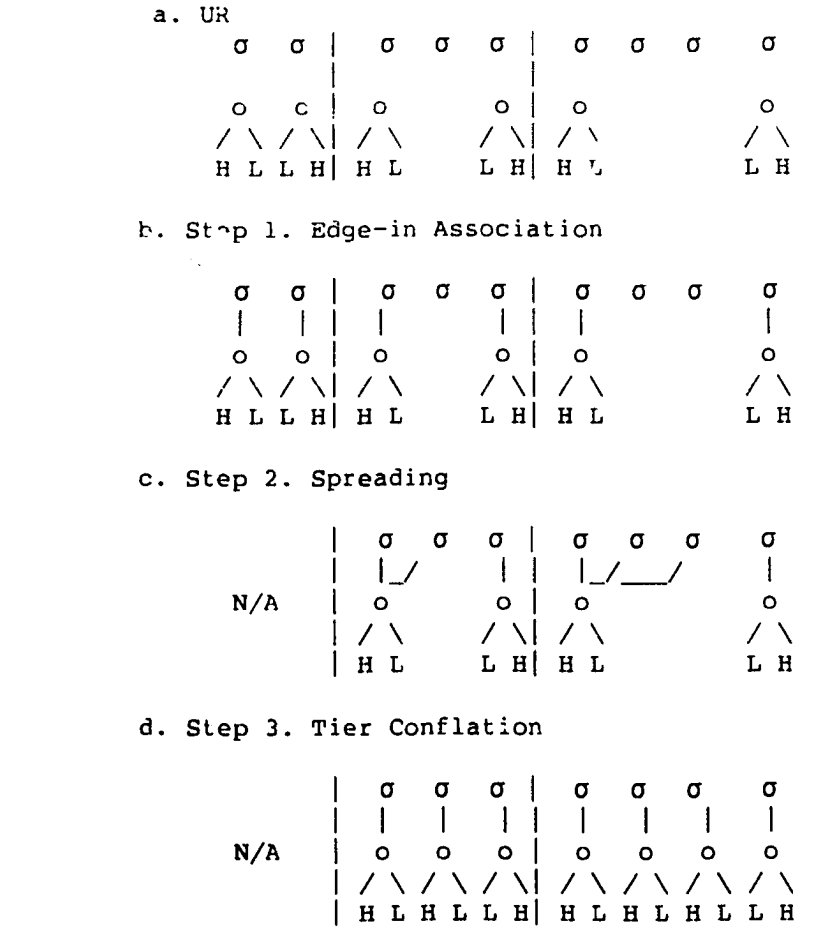
Danyang is a dialect of Wu, one of the Chinese languages that centre around Shanghai. In the data from Danyang, originally cited in Lü (1980), Yip finds that contour tones may assimilate (or spread) as units. The following is one example from the data[8]:

(2.32).

dissyllabic word	trisyllabic word	quadrisyllabic word
42-24	42-42-24	42-42-42-24

What is illustrated above is a word tone in Danyang. One characteristic of such a word tone is that its contour remains constant at the two edges of the domain of a word, no matter how long the domain is. Meanwhile, any syllables that occur in between the edges will take the tone of the first syllable. Yip contends that "this kind of repeated contours is only explicable by spreading or copying the contour of the first syllable as a unit (Yip 1989b, p.162)." Representing (42) and (24) as HL and LH respectively, Yip proposes the following analysis for the above example:

(2.33). 42-24 42-42-24 42-42-42-24



where N/A = non applicable

First, there is the edge-in association through which the initial and final syllables of a multi-syllable word are associated with the contour-tone nodes also at the two edges of the word. Then, left-to-right spreading occurs thereby the medial syllables get associated with the first tonal node of the word. The last step of tier conflation yields the actual surface representation of these syllables and tones.[9]

This analysis of the Danyang case clearly shows: one, the edge syllables associate (Step 1) with the edge contours as units (i.e. the association involves the tonal node "o" rather than the tonemes). Two, the medial syllables are associated (Step 2) to the first tone as a unit (i.e. the initial contour tone (42/HL) rather than

any of its elements spreads as a unit). Such a case of Danyang word tone, as Yip points out, cannot be reasonably analyzed if contour tones are not allowed to associate as units.

2.8.2 Evidence from Tianjin

Other evidence provided by Yip to argue for the existence of (2.30a) is that "the OCP[10] can identify complete contour tones as units, and then trigger a rule of dissimilation (Yip 1989b, p.163)." Her evidence is from Tianjin tone-sandhi data some of which are illustrated below:

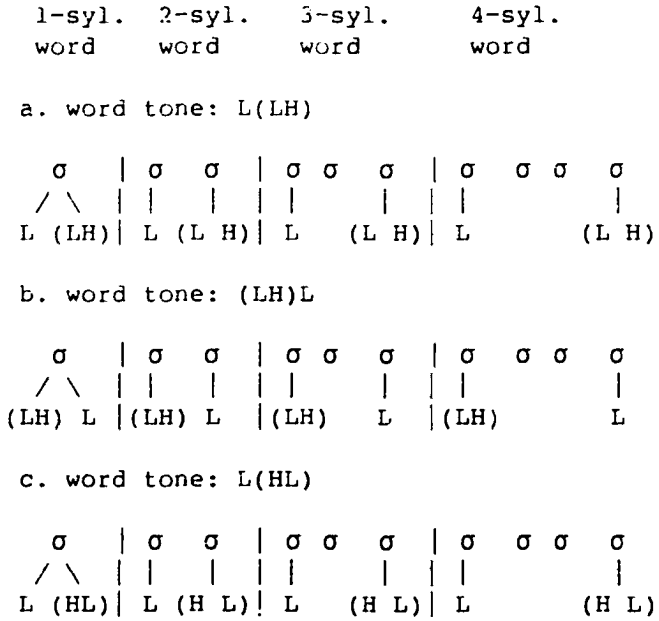
(2.34). a. LH.LH → H.LH but, d. LH.HL remains LH.HL
b. HL.HL → L.HL e. HL.LH remains HL.LH
c. L.L → LH.L

In the above, the cases from (a) to (c) clearly demonstrate that the identity in the two adjacent tones triggers the operation of dissimilation so as to comply with the universal OCP. Similar sandhi, on the other hand, has not occured to non-identical contour tones even though they contain identical tonemes adjacent to each other at the tonal boundaries (d and e). Such a dissimilation process triggered by the OCP would be inexplicable, argues Yip, if the unitary structure of the contour tones were not recognized.

2.8.3 Evidence from Wuxi

Tone melody in Wuxi, another dialect of Wu, constitutes the most important evidence given by Yip for the branching type of analysis (2.30a). According to Yip, Wuxi has the following four word tone melodies: L(LH), (LH)L, L(HL) and H, each of which is found to flank both sides of a domain of a word. Here I will discuss the three more straight-forward melodies among them only[11].

(2.35).



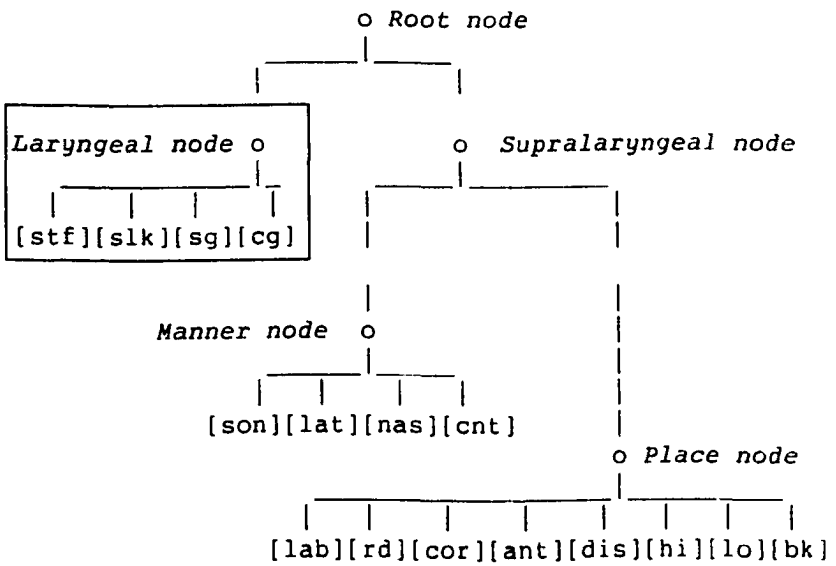
One thing that is shown above is that the four word tones are invariably "squashed up or stretched out to fit the domain (Yip, 1989a, p.43)."[12] In the process of the "squashing up or stretching out", the contour elements of the word tone, LH in (a) and (b) and HL in (c), always associate as a unit no matter how long the domain is. Thus, the contour tones in Wuxi, argues Yip, must be allowed to operate as units so as to maintain the above simple and elegant analysis. Otherwise, it is not clear how the patterns in (2.35) can be captured in a principled way.

I think that Yip's arguments, as described above, for the inclusion of contour tones as units in the underlying representation are quite convincing. It is clear that any phonologist who should reject the contour-tone-as-unit analysis would have great difficulty in handling these Chinese cases.

2.9 *Bao (1990)*

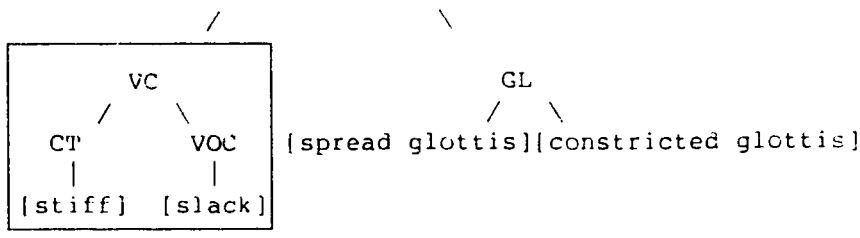
After Yip (1980a), the next major comprehensive study of Chinese tonal representation is found in Bao (1990). Just as Yip's (1980a) study reflected the theoretical advance in generative phonology made in the late 1970s, Bao's reflected the trend of the late 1980s in the field. The trend of generative phonology has, since the middle 1980s, turned to the non-linear configuration of classes of features, an area of study under the name *feature geometry* (cf. Clements 1985, 1989, Goad 1991, and McCarthy 1988). The result of the research in this area consists of a tree which marks the relationship among features and natural groups of features. Various shapes of the tree have been put forth,[13] an example of which is provided below (from McCarthy 1988, p. 89):

(2.36).



One of the purposes of Bao's study is to work out in more details the laryngeal node shown in the box. The following is what he proposes for it:

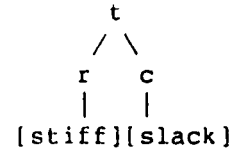
(2.37).



where
VC= vocal cords
GL= glottis
CT= cricothyroid
VOC= vocalis

The major concern in Bao's study is the VC node in the geometry above. This node, according to Bao, is the node for tone, and the local tree in which it is the mother corresponds to the following local tree.

(2.38).



where t=tone
r=register
c=contour

Bao gives the following explanation for their correspondence:

> The geometry of tone is the geometry of the VOCAL CORDS, which I take to be the articulator that executes tone. It is speculated that the cricothyroid executes the register feature [stiff] and the vocalis executes the contour feature [slack]." (pp. 10-11)

With the tone geometry in (2.38), Bao stipulates that a tone consists of two components, register and contour. This stipulation, as one may recall, is nothing new: the idea of a register node is seen earlier in Yip (1980a, 1989a,b). This above geometry also shows that each of the sister nodes dominates one binary feature, [stiff] for the register node, while [slack] for the contour node. According to Bao,

the contour node, or the c node, can optionally branch, although it is confined to no more than a binary branch by the following constraint:

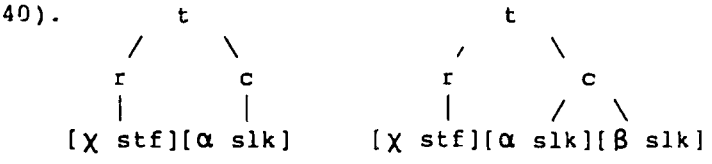
(2.39).

> "Underlyingly, the contour node may have at most two branches." (p.61)

The register (or the r) node, on the other hand, is not permitted to branch although somehow no explicit rule is stated in Bao to block the branching. Such a treatment of the register node, as one may recall, serves the same purpose as Yip's constraint which requires her Register Feature [upper] to remain constant in value within the same morpheme (cf. 2.2-).

Given the explicit and implicit constraints, the power of the tree in (2.38) is reduced to the generation of the following types of trees only.

(2.40).



where χ = either α or β

As previously implied, Bao is fairly similar to Yip (1989a,b). In fact, he is identical to Yip in several respects. Just as Bao himself put it, the two features "[stiff] and [slack] are functionally equivalent to [upper] and [high] of Yip (p.59)". By "functionally equivalent", Bao means that his two features define the same four tonal levels as defined by Yip's features, and that these levels are defined in the same way as Yip's four levels of tones are defined. The following is a more visual comparison of the two definitions:

(2.41). Yip (1980a, 1989a,b) Bao (1990)

	[+high]			[-slack]
[+upper]			[+stiff]	
	[-high]			[+slack]
	[+high]			[-slack]
[-upper]			[-stiff]	
	[-high]			[+slack]

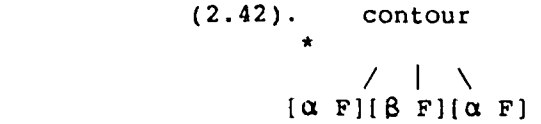
Bao justifies the use of his two features rather than Yip's by pointing out that Yip's features indicate auditory impressions while his features, adopted from Halle and Stevens (1971), show articulatory considerations. He observes that "the advantage of Halle-Stevens' feature system is its ability to express the pitch of vowels and voicing of consonants as featurally the same phenomenon (p.57)."

While it may be true that Bao's features better represent the articulatory gestures, I have found that the difference between Yip's and Bao's features is trivial: it is but a matter of difference in the preference of terms rather than of any profound theoretical differences. At least, as far as the phonological analysis of the Chinese tones is concerned, Bao's features are identical to Yip's.

Besides being identical to Yip in the use of the tonal features, Bao has another feature in common with Yip: he is a supporter of Yip's (1989a,b) hypothesis that contour tones may act like atomic units and, therefore, should be represented as units at some level of the representation. This position of Bao's is clearly indicated by his use of an individual contour node (i.e. the c node) in his trees in (2.40).[14]

Another similarity between Bao and Yip is that both allow for binary-only branching of the contour node. Such a treatment implies that both regard the following tree as unacceptable:

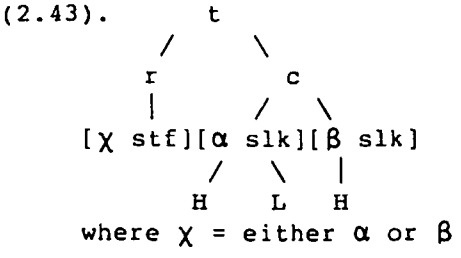
(2.42).



where
F = a variant for any feature

Of course, besides the similarities, Bao also differs from Yip in several ways. First, a difference exists between them in terms of whether complex contour tones should be represented in deep structure. This question is tantamount to asking, in technical terms, whether any of the two branches of the contour node (or the c node) should be allowed to branch further; or whether a tree of the following shape is allowed:

(2.43).



To Yip, the answer is yes, and thus she permits further branching of one (and only one) of the two branches and allows for a sequence of three terminal nodes to occur underlyingly. Bao, on the other hand, takes no for the answer, and thus does not allow such a branching at all. This position of Bao's is obvious in his constraint mentioned earlier in (2.39). According to Bao, his constraint is to[15]

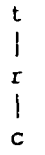
(2.44).

> "imply the proposition [that] concave/convex tones are surface phenomena." (p.62)

By now, one may have already discovered a major difference between Bao and Yip. To facilitate the understanding of the difference, let us first recall the two unique characteristics that distinguish Bao and Yip from their predecessors. The two characteristics are: one, both Bao and Yip are in favor of treating the contour tones as units as well as sequences of level tonemes in the underlying representation. Two, both employ the notion of *register* to encode the traditional two-way *yin* and *yang*[16] partition of the pitch range. However, Bao and Yip differ in terms of the structural relation between the register and the contour node. In Yip, the

register node and the contour node are co-extensive in the sense that the former dominates the latter. In Bao, however, they stand as sisters to each other. If Bao's node names are used for Yip's tree, the two can be contrasted as follows:

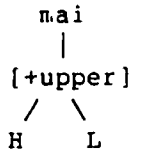
(2.45). a. Yip (1989a,b)



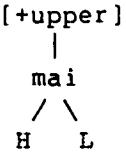
b. Bao (1990)

In fact, Bao's treating the register and the contour nodes as sisters to each other is not totally new. It is the same treatment seen in Yip's 1980 dissertation. For example, the sister relation is indicated in her representation of the syllable *mài* in (b) in the following. This representation is different from the dominance relation between register and contour implied in her later studies shown in (a) in the following. Notice the structural equivalence between (b) below and (b) in the above (2.45).

(2.46). a.



b.



2.10 Conclusion

There is no doubt that the studies reviewed in this chapter have made a tremendous contribution to our present understanding of the tonal phenomena in languages, particularly in Chinese. However, it is felt that further in-depth studies of the individual languages covered in these studies are urgently needed if one wishes to substantiate any of these theories, their tenets, claims and assumptions.

2.11 Notes to Chapter 2

1 This is a published version of the author's 1976 dissertation.

2 See Anderson (1978) for details of these studies by Gruber, Sampson, and Maddieson.

3 Woo (1972) is a published version of the author's 1969 MIT dissertation.

4 The data given below are in Woo's original transcriptions.

5 Goldsmith's arguments have been cited in many studies. I therefore will not repeat them here. For details of these arguments, the reader is referred to Goldsmith himself.

6 The constraint should really be stated as being within the domain of the syllable rather than the morpheme. See Chapter 3 for more discussion of the problem.

7 This argument is apparently complementary to a reverse statement that other languages have contour tones that do not behave as units. This latter argument has no great controversy surrounding it, though.

8 The Danyang data are also discussed later in § 7.3.7.

9 One of the main objectives in Yip's citing of the Danyang data is to argue for a unique type of association which links the units of the two tiers at the two edges of the relevant domain first before moving inwards. Before Yip, there were only two types of association observed: left to right or right to left. Compared with them, Yip's is a bi-directional type of association.

10 The letters OCP stand for Obligatory Contour Principle. Roughly speaking, OCP is a principle whose basic function is to dissimilate: it prohibits adjacent identical segments. This principle was first formulated in Leben (1973) and Goldsmith (1979) specifically for tonal phenomena, but has since been extended to segments of other features, and has been formalized further in details with regards to various features at various levels (e.g. McCarthy 1986, 1988, Yip 1988, and Odden 1988). It has been widely cited as follows,

> "At the melodic level, adjacent identical elements are prohibited." (McCarthy 1986, p.208)

and has been observed to perform a bewildering number of functions as a morpheme structure constraint, rule blocker, rule trigger, constraint on the mode of operation of an ambiguous rule, and constraint on the form of possible rules (Yip 1988).

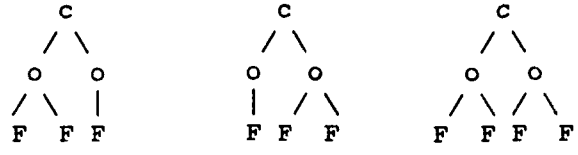
11 The Wuxi tones are also discussed in § 7.3.8.

12 I will ignore the tonal instantiation on the medial syllables as it is of no concern here.

13 Due to the complexity of the theory of feature geometry, I am not providing a review of the various frameworks proposed so far.

14 Bao does not explicitly claim that he is a follower of Yip in this respect. It is not stated in Bao whether he has independently worked out the idea, or is merely supporting Yip's. Given the time span between these studies, either situation is possible.

15 Here, a technical problem in Bao should be pointed out: his constraint in (2.39) does not imply (2.44) as it is meant to. In fact, nothing in it would prevent the occurrence of trees such as the following,



Nothing in Bao would prevent the two original branches from branching further and resulting in a sequence of more than two terminal nodes. This is clearly shown above, where none of the c nodes has more than two immediate branches, which means that none violates Bao's constraint in (2.39). In fact, even trees with more levels than those in the above illustration (i.e. with more than three terminal nodes), would still observe the constraint in question.

It is quite obvious that, for the constraint to work as desired, Bao needs to

add to it another phrase such as "..immediately dominating two and only two terminal nodes." Despite the technical problem just mentioned, Bao's position of no complex contour tones in the underlying structure is however clearly expressed.

16 In Chinese historical phonology, there is a well-attested theory that each of the four Middle Chinese (around the 6th century) tones was split into two during the period between Middle Chinese and the early vernaculars of the 13th century. The split was originally conditioned by the difference in the voicing of the initial consonants of the syllable. The voiced initial is said to condition a higher pitch, whereas the voiceless one a lower pitch. Later, most of the Chinese languages experienced a loss of the voiced initials through a merging process in which all the voiced initials merged with their voiceless counterparts. After the loss, however, the conditioned pitch variance remained. The result was that what had before been conditioned phonetic pitch difference became now contrastive and hence Chinese became a eight-tone (rather than a four-tone) language. These eight tones are four higher pitch ones called *yin* tones and four lower pitched ones called *yang* tones.

According to this theory, the various tonal systems seen today in the various Chinese languages have evolved (largely through a merging process) from these eight *yin* and *yang* tones. For more discussion of this historical development, please refer to § 3.5.

Chapter III

THE REGISTER FEATURE SYSTEM

In this chapter, I will discuss in detail Yip's (1980a) system of tonal features, particularly as it applies to Mandarin tones. Yip's system with its unconventional Register Feature has exerted a major influence on subsequent studies of Chinese tones. Being virtually the only non-linear analysis of Chinese tones for more than a decade, Yip's two-layer system has been adopted in a number of studies of Chinese tones (e.g. Bao 1990, Shih 1986, Packard 1989, Pulleyblank 1986, Yip 1989a,b, among others). These studies, however, have built further theories on the basis of Yip's system without seriously questioning its fundamental validity. In fact, few studies so far have seriously challenged it. One notable exception is a recent study by Zee (1991), who argues that Shanghai is fundamentally a three, rather than four level tonal system (H, M, L), and hence that Yip's Register Feature system cannot be adopted for the analysis of Shanghai tones.[1] As I will show below, my review has also found Yip's analysis (in particular, her Register Feature) unmotivated for Mandarin.

Thus, two claims may perhaps be made regarding Yip's analysis, a strong one and a weak one. The strong one claims that Yip's Register Feature is inadequate in the analysis of Chinese tones in general, and the weak one asserts that the feature is irrelevant in the analysis of Mandarin tones. In the present study, however, I will try to make the weak claim only, leaving the proof of the strong claim to later studies. The reason for this narrower focus is that the task of proving the strong claim goes far beyond the scope of this study. Regarding Chinese languag-

es other than Mandarin, I will simply say that if Yip's Register Feature can be proven applicable in any of them, it is likely a language specific feature, for at the very least, it cannot be maintained for Mandarin.

Before beginning the discussion of Yip, let us refresh our memory of her analysis, briefly introduced in the last chapter. The most unusual aspect of Yip's Register Feature system is probably its use of an unconventional feature *[upper]*. This feature and the feature *[high]* are the two component features in Yip's analysis. Called Register Feature and Tonal Feature respectively, they interact to define four tonal levels:

(3.1).

Register Feature	Tonal Feature	Chao's Scale	Four Tonal Level Defined
[+upper]	[+high] (H)	5	[+upper, H]
	[-high] (L)	4	[+upper, L]
[-upper]	[+high] (H)	2	[-upper, H]
	[-high] (L)	1	[-upper, L]

Claiming a separate autosegmental tier, the Register Feature partitions the whole range of the voice pitch into two equal portions, a [+upper] portion and a [-upper] one. On Chao's five-level scale, [+upper] covers a pitch range from 5 to 3, and [-upper] from 3 to 1. Each of the two portions, in turn, is divided into two sub-portions by the Tonal Feature [high]. Thus, in terms of coverage of the pitch range, these two features overlap each other. Because of this overlapping nature, I may perhaps refer to Yip's analysis as a two-layer system to distinguish it from other one-layer systems whose features do not overlap this thoroughly in pitch coverage.

In all, Yip cites four pieces of evidence from Mandarin to motivate her Register Feature analysis. I will go through them one by one below.

3.1 *Mandarin Case 1: the Neutral Tone*

First, let me give a brief introduction to the general characteristics of the neutral tone in Mandarin. In the literature, a syllable with the Mandarin neutral tone has been described as toneless or atonic (Dow 1972), or pitchless (Li and Thompson 1981). It is said to be short and light (Xu 1983), and short and lax (Cheng 1973). Based on earlier acoustic studies by Zadoenko (1958) and Dreher and Lee (1966), Cheng concludes that "the length of a neutral tone syllable is about one half that of a full tone syllable (p. 55)."

The neutral tone has been observed to be related to no stress or lack of stress. Being the first person using the term "neutral tone" for the tone in question, Chao (1968), for instance, indicates his belief in a connection between the two by placing the discussion of the neutral tone in the section entitled "Weak Stress" in his book. In this study, Chao also remarks that "in weak stress, the tone range is flattened to practically zero and the duration is relatively short (p. 35)."

The connection between the neutral tone and lack of stress is also expressed in Li and Thompson (1981). According to these authors, "if a syllable has a weak stress or is unstressed, it loses its contrastive, relative pitch and therefore does not have one of the four tones. In such a case, the syllable is said to have a neutral tone (p. 9)." In still another study, the neutral tone and weak stress are not even distinguished. Xu writes in his 1983 paper that "the neutral tone is also called weak stress. It is pronounced with the characteristics of being light and short (p. 220).[2]" To establish a cause-effect relationship between the neutral tone and stress, Cheng (1973) maintains that the neutral tone is derived from lack of stress. According to him, "when a syllable is stressed, it has a tone, but when it is unstressed, its tone becomes neutral (p.57)." Therefore, he continues, "the neutral tone items must be specified with full tones in the lexicon (p. 66).[3]"

Still another fact concerning the neutral tone is that almost all morphemes that appear in the neutral-tone shape have also a corresponding full-toned form found in a stressed position. This fact provides indication that the neutral tones in these morphemes are each derived from their respective full tones in a non-stressed or weak-stressed situation. There are, however, a handful of exceptions (fewer than ten), and these include the frequent grammatical particles such as the perfective particle *le*, the possessive particle *de*, and the continuous-aspect particle *zhe*.[4] These particles are always in the neutral tone form, and their exact full-tone correspondences are hard to trace from a synchronic point of view.[5]

What then is the shape of the neutral tone in phonetic or phonological terms? According to Chao (1968), the neutral tone does not have a phonemic tone shape (or pitch value) of its own; rather, it derives its pitch value from tones that precede it. When following other tones, for instance, its pitch value varies with the pitch values of the end points of the preceding tones.

Having given a general introduction to the neutral tone, let us examine Yip's analysis of it. Yip indicates that the neutral tone (Tone 0) provides evidence for the autonomous status of the [upper] feature, and therefore justification for its postulation. Specifically, she argues that this Mandarin tone, when appearing on the handful of grammatical particles just mentioned, is prespecified for its Register Feature [upper], but not for its Tonal Feature [high]. Namely, it consists of the following underlying representation:

(3.2). [-upper]

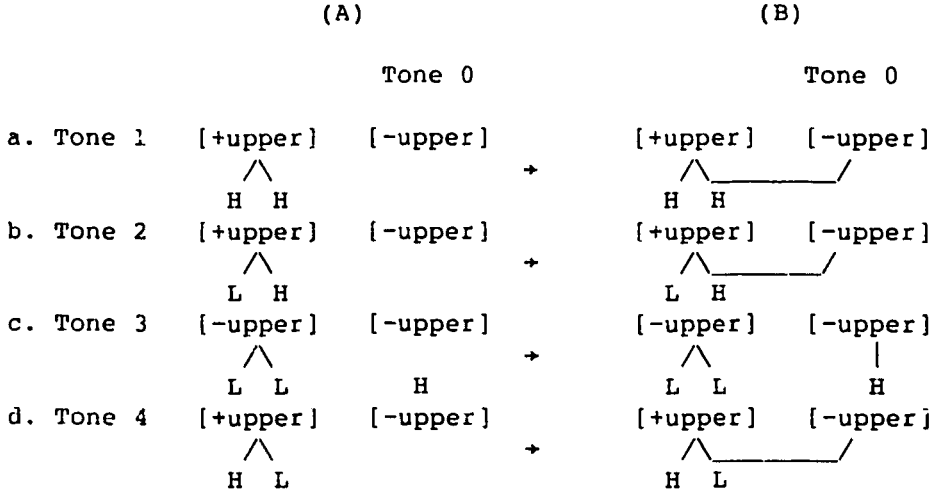
Yip's analysis for the neutral tone is based on the following data (Yip, p.47 and p. 163).[6]

(3.3).

			Tone 0	Examples		
Tone 1	(55)	3		chī	le	"have eaten"
Tone 2	(35)	3		ná	le	"have taken"
Tone 3	(21)	4		mǎi	le	"have bought"
Tone 4	(53)	1		huài	le	"have gone rotten"

These data indicate specifically that the neutral tone has a middle-level pitch after Tone 1 and Tone 2, and a low pitch after Tone 4, but a relatively high pitch after Tone 3. If Chao is right in considering that the value of the neutral tone is derived from its preceding tone, what then are the derivation processes involved? According to Yip, the derivations are done in the following manner:

(3.4).



First, in each case, the neutral-toned morpheme shows up with its prespecified feature [-upper] after the concatenation of the two relevant morphemes. Then, spreading occurs through which the neutral tone acquires the value of its Tonal Feature from its preceding tone. Such an analysis works well in the cases of Tones 1, 2 and 4, but fails to work for the neutral tone following Tone 3. Yip's solution for this problem is "a special rule that inserts a H tone after the third tone when no other tone follows (Yip, p. 162)." Yip explains that by "no other tone follows", she means "pre-pausally or before a neutral tone (p.162)". Namely, a rule in the following shape:

(3.5). LL → LLH / ___ {pause, a neutral tone}

Yip's explanation is unconvincing. While it seems quite natural to consider a pause as a case of "no other tone follows", it is not exactly clear whether the neutral tone can be similarly considered. It is not clear how a syllable, even in the neutral tone, can function just like a pause. Arguably, the appearance of an extra toneme in the third tone before a pause may be due to the fact that there is "room" for its appearance. The pre-pausal position is after all a common place where extrametrical and idiosyncratic materials do occur to "take shelter" from rigid phonological rules. No such "room" or shelter-type function, however, can be found before another syllable, neutral-toned or otherwise.

Or, from another viewpoint, if the occurrence of an extra toneme is to achieve a growth in length in the third tone so that it may realize its potential full-length, no such tendency toward a length-increase is seen in the pre-neutral-tone position. As a matter of fact, a research finding made in Dreher and Lee (1966) shows that the tone before the neutral one is shorter than it is before other tones. According to these authors, tones immediately preceding the neutral tone are about 20 percent shorter than usual.

The conclusion to be drawn from the above discussion is that there is no bona fide reason for the insertion of a H toneme in that specific position (see 3.4c). It should be noted, however, that this problem in Yip's analysis does not constitute a sufficient argument to reject her analysis. That argument comes from the fact that Yip's analysis of the neutral tone fails to capture an important generalization about the neutral tone. In fact, this failure (which I will show later) does not originate from her own work, but it stems from an erroneous understanding of the data that her analysis is based on.

As mentioned earlier, Chao believes that the neutral tone does not have its own inherent pitch value, but acquires its surface value from its preceding tone. Chao's observation is true to a certain extent. The proposed assimilation (or spreading, shown in B in 3.4) does indeed occur to produce the precise surface pitch forms. The question is: how can one account for these precise pitch values? Or, is there any generalization hidden among these phonetic surface values?

The answer to the latter question is yes. A more detailed discussion and justification of the existence of this generalization is provided in § 6.7; I will give only a rather brief demonstration of it here. Now, let us have another look at Yip's data provided earlier in (3.3).

(3.6)

	Tone 0
Tone 1 (55)	3
Tone 2 (35)	3
Tone 3 (21)	4
Tone 4 (53)	1

One observation that can be made about these data is that the differences in pitch value between the endpoints of Tones 1, 2 and 4 and that of the following neutral tone all equal 2: in all three cases, the pitch of the neutral tone is 2 points lower than the endpoint of its preceding tone, measured on Chao's scale. However, the difference in value between the neutral tone and its preceding Tone 3 is a negative 3, (if the number 2 just mentioned is regarded as being positive). The question is: why does the neutral tone behave differently when following the third tone?

The answer clearly lies in the distinct form of the third tone as opposed to the forms of the other three phonemic tones: it is the only low tone and the only tone that reaches the lowest pitch before the neutral tone. This fact does not seem to matter much when one tries to relate it to the precise phonetic values of the neutral tone. But suppose the neutral tone is represented as a fully specified low tone

in the underlying representation, the problem would become self-evident. This can actually be illustrated in Yip's analysis.

(3.7).

		Tone 0
a. Tone 1	[+upper]	[-upper]
	/\	|
	H H	L
b. Tone 2	[+upper]	[-upper]
	/\	|
	L H	L
c. Tone 3	[-upper]	[-upper]
	/\	|
	L L	L
d. Tone 4	[+upper]	[-upper]
	/\	|
	H L	L

One thing that becomes readily observable from such a configuration is the total identity between the shape of the neutral tone and the endpoint of its preceding Tone 3 (c). Notice that the same identity is not found in the other cases. Under such an analysis, the mystery about the odd behavior of the third tone is resolved: the identity triggers a dissimilation process which raises the neutral tone after Tone 3 in pitch. Namely,

(3.7c'). Tone 3 [-upper] [-upper]

/\ \|

L L L → H

With this above generalization captured, it becomes now easy to account for the other values of the neutral tone. Clearly, the higher surface values of the neutral tone after Tone 1 and Tone 2 are due to a lower level phonetic coarticulation, as described in Shen (1990b). Specifically, the value of the neutral tone is raised to 3 simply because Tones 1 and 2 that precede it are both tones ending in high tonemes. It is not thus after Tone 4 because Tone 4 ends in a relatively lower toneme.[7]

It should be noted that it takes the identity in both tiers to trigger the dissimilation process, described in (3.7'), as a similar process is not found in the case of

Tone 4 where there is no match between the two Register-Feature specifications in spite of the fact that there is identity at the Tonal tier. If this observation is correct, it should serve to indicate that the neutral tone cannot be an underspecified tone, but just a plain low tone (probably 1 in pitch value), and that the neutral tone case does not really constitute an argument for the autonomous behavior of the proposed Register Feature.

3.2 Mandarin Case 2: Tone on Post-Verb Syllables

Another similar piece of Mandarin evidence Yip uses to motivate the Register Feature [upper] and its autosegmental behavior comes from the following data (Yip, p. 63 & p. 175):

(3.8).

mài	"sell"	sòng	"deliver"
shàng	"up"	qù	'go, to"
le	"perfective particle"		

a. mài le. "sold; have been sold"

b. sòng shàng qù "deliver up (to)"

c. sòng shàng qù le "delivered up (to)"

One characteristic concerning the above three sentences is that all of the syllables after the main verbs *mài* and *sòng* receive weak stress, and the stress pattern can be roughly shown as follows:

(3.9). a. s w
mài le. "sold; have been sold"

b. s w w
sòng shàng qù "deliver up(to)"

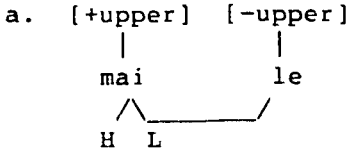
c. s w w w
sòng shàng qù le "delivered up(to)"

where s=strong, and w=weak

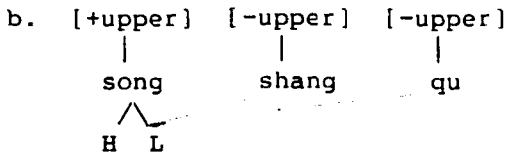
Namely, the main verb in each case is stressed while the rest of the syllables are not (or have weak stress). In addition, these post-verb syllables may be regarded

as carrying the neutral tone,[8] although the source of the neutral tone on *le* is different from that of the neutral tones on the rest of the post verb syllables. The former is underlyingly a neutral tone (signalled by the absence of any tonal diacritic); it is a neutral tone before syntactic concatenation. The latter, however, are underlyingly full-toned, derived through a lack of syntactic stress. In general, all the syllables after the verb are perceived as rather low in tonal value. Now let us see how Yip interprets the data as a support for the postulation of the Register Feature:

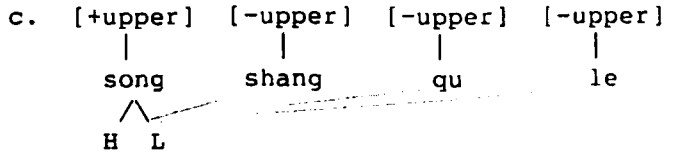
(3.10). a.



b.



c.



Assuming that all the neutral toned syllables are prespecified as [-upper] for the Register Feature, Yip adopts a spreading analysis whereby all the post-verb syllables acquire their low tonal value through the spreading to them of the Tonal Feature L ([-high]) from the preceding verb.

First, there is a minor problem regarding this analysis: while Yip is explicit about the source of the feature [-upper] on the syllable *le* (one of those few syllables in Mandarin which are always neutral-toned; see also the last section for relevant details), she does not explain how [-upper] gets there for the rest of the post-verb syllables. However, this problem is probably not difficult to work out; it

appears to be just a technical problem which needs to be worked out in detail. Therefore, I will dwell no further on it.

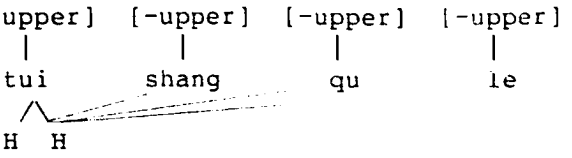
The major problem with this analysis lies in its violation of the locality requirement implicit in the observation that the neutral tone derives its tonal value from its preceding tone. Although the neutral tone has been reported to derive its phonetic surface value through spreading from its immediately preceding tone, no studies have found this spreading to go beyond its immediately following tone. This problem does not show much in Yip's analysis (3.10) of the sentences in (3.8), since this analysis does yield the grammatical result of a low tone on these post-verb syllables. The problem is better shown if Yip's data are expanded to include the following sentences:

(3.11). tái "carry"
tuī "push"

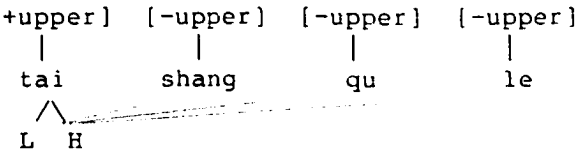
a. tūi shang qu le "pushed up"
b. tái shang qu le "carried up"

Notice that in Yip's earlier data, all the main verbs are by chance in the fourth tone (51). These two sentences here are identical to her sentence in (3.8c) in every other aspects except for the tonal values of the main verbs at the beginning of the sentences. Here in (a), the verb *tui* carries the first tone (55), while in (b), the verb *tai* is in the second tone (35), both being tones ending in high tonemes. Based on Yip's earlier treatment, the neutral-toned, post-verb syllables in these two sentences should acquire their tonal value through prespecification and then spreading as shown below:

(3.12).* a.

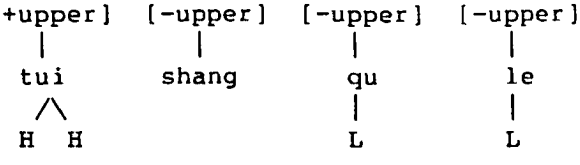


* b.

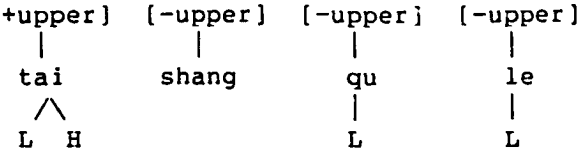


While the results in (3.10) may by chance be grammatical, the results here with all the neutral-toned syllables to have a value at the middle range (i.e [-upper, H]) are clearly unacceptable. Although this value (i.e [-upper, H]) may be accepted for the syllable IMMEDIATELY following the verb as a low-level phonetic surface representation, it can by no means be accepted for the two tones that follow. The right output for these two tones should be just as the post-verb tones in the earlier cases in (3.10):

(3.13). a.



b.

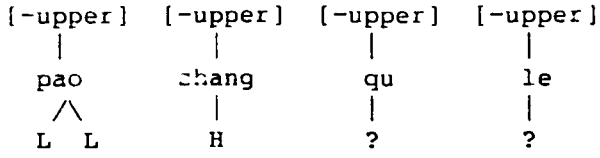


The problem involved is this: the spreading will not go beyond its adjacent tone, if there is indeed a spreading of the Tonal Feature from the preceding tone. From this follows another problem: if the two final syllables do not receive their Tonal Feature specification from the main verb, where can they acquire that feature specification? Or, how exactly can they surface fully specified? It is not clear how these questions can be answered in Yip's analysis. The problem of the source of the Tonal Feature specification shows up even more obviously in another sentence of the same type:

(3.14). pǎo "run"
pǎo shang qu le "ran up"

This sentence is once again identical to Yip's sentence in (3.8c) in every respect except for the tonal value of the main verb. The difference is that there, the verb is fourth-toned (51), but here the verb is third-toned (213). Now a serious problem arises in the acquisition of the Tonal Feature specification on the part of the post-verb syllables. Let us observe the following:

(3.15).



Recall that in Yip, the neutral tone immediately after the third tone acquires its Tonal Feature specification by way of the following rule (cf. (3.5)).

(3.16) LL → LLH / ___ {pause, a neutral tone}

What this rule says is that a Tone 3 (213), represented as [-upper] LL, acquires a H tone when preceding a neutral tone or a pause. If the locality requirement on the spreading of the tonal value from a preceding tone to its adjacent following neutral tone is not obvious in the previous cases in (3.8) and (3.11), it should be quite explicit by virtue of this rule that the only neutral tone that is affected is the one IMMEDIATELY after the verb. However, although this rule is right in correctly encoding the adjacency requirement, it also reveals a problem in Yip's treatment of the type of sentences in question. The problem is: how can one account for the tonal values of the two syllables that are not adjacent to the verb? It is not clear how this difficulty can be easily surmounted. It would seem rather unnatural if more rules should be devised to insert more H tones for the two words *qu* and *le* at the end of the sentence. Even if these H tones can, by

some perhaps unnatural means, be created for the two syllables, the resultant sentence is ungrammatical anyway. This is because the last two syllables again do not occur as a middle tone [-upper, H], but rather as a low tone, specifiable in terms of [-upper, L].

By now, it should be clear that Yip's analysis for the post-verb tones under discussion is not adequate, and that it has not succeeded in demonstrating that the autonomous Register Feature is supported through this analysis. Incidentally, this case of the tones on the post-verb syllables provides further evidence to show the implausibility of Yip's analysis for the neutral tone discussed in the last section. Among other problems, it cannot account for the surface tonal shape of a neutral-toned syllable found after another neutral tone. And, this case provides further evidence that the neutral tone should be fully specified before the low-level assimilation takes place.

3.3 *Mandarin Case 3: the "Conclusion" Intonation*

Another piece of evidence Yip draws on to support her positing of the Register Feature comes from sentences with "conclusion intonation" first described in Chao (1968). The following are examples of these sentences:

(3.17). Clause 1 Clause 2

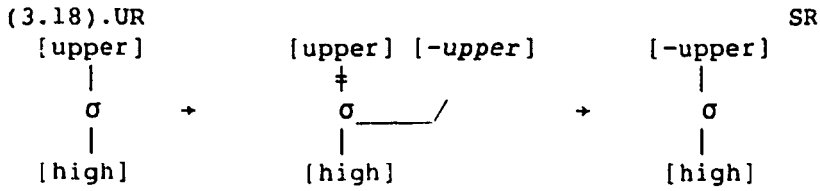
a. nǐ xìng Wáng, wǒ xìng Lù.
"You have the surname of Wang; I have the surname of Lu."

b. wǒ xìng Lù, nǐ xìng Wáng.
"I have the surname of Lu; You have the surname of Wang."

Certain features apparent in these sentences are: one, each contains two clauses -- let me label the first clause in each sentence Clause 1 and the second in each Clause 2. Two, Clause 1, according to Chao (1968, p.39-40), ends in a rising intonation, while Clause 2 ends in a falling intonation. Three, the rising intonation

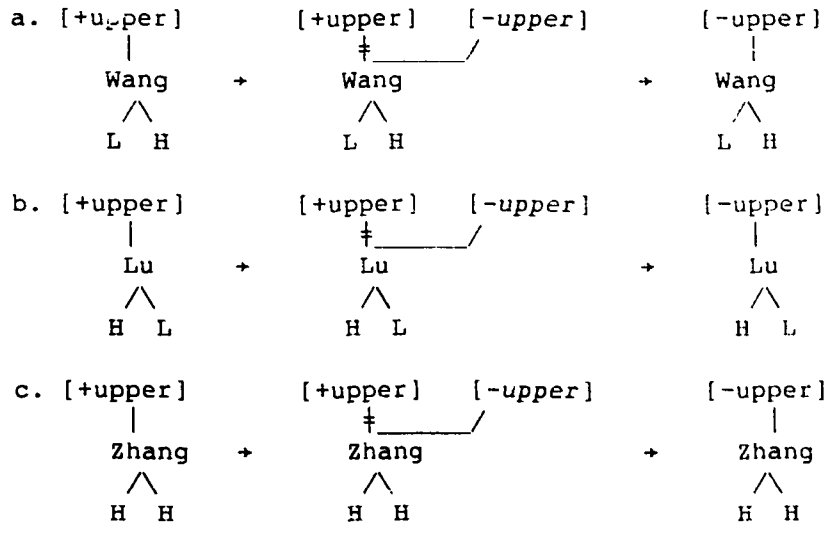
of Clause 1 causes the rising second tone (35) on *Wang* (a) and the falling fourth tone (51) on *Lu* (b) to raise their pitch higher as a whole while still keeping their original tonal contour. The falling intonation of Clause 2 causes the final syllables *Lu* in (a) and *Wang* in (b) to be pronounced with a generally lower pitch, but again, still keeping their own original tonal contour.

Yip's analysis of Chao's two sentences covers the falling intonation at the end of Clause 2 only, without saying anything about the rising intonation on Clause 1. Specifically, Yip attributes the dropping of the tonal value on the final tone of Clause 2 to the presence of a [-upper] Register Feature floating at the end of these sentences. This floating Register value (i.e. [-upper]) then overrides the underlying Register value of the tone occurring in this specific sentence-final position. As a result, the syllable surfaces with a lower Register value. Yip's analysis is illustrated below: (The overriding floating Register Feature is italicised.)



What her analysis says is that a Mandarin tone of any shape occurring in the sentence final position in question will delink from its own underlying register tier and link to the register feature floating in that position. Thus, when the morpheme occurring in that position has a Tone 4 (e.g. *Lù*) or a Tone 2 (e.g. *Wáng*), or to add another common Chinese surname to the data, a Tone 1 (e.g. *Zhāng*), the application of this above rule will yield the right surface representations, as shown below in (a) and (b):

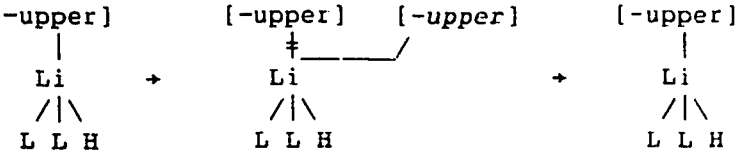
(3.19). UR SR



The result is that the syllables *Lù* , *Wáng* and *Zhāng* , all of which are underlyingly [+upper], come out of the derivation with a desired [-upper] specification to signal that the falling intonation has occurred on them. So far, Yip's analysis seems quite successful in deriving and representing the lowered pitch value, the result of the falling intonation. And her analysis will succeed as long as the underlying register value is [+upper].

However, a problem arises when the tone in the underlying representation is [-upper].

(3.20). [-upper] [-upper] [-upper] [-upper]



In the above, *Lǐ* is another common Chinese surname, which has a Tone 3. Now let us see how Yip's treatment works if this tone also appears in the position of the falling intonation. In Yip, Tone 3 is the only [-upper] tone of the four Mandarin phonemic tones. Being already specified [-upper], Yip's analysis applies vacuously to this tone to yield an output that is identical to the input. The identity of the

SRs and URs means that the analysis has not encoded the falling intonation in this case of the Third Tone. It thus seems that the success of Yip's analysis in (3.18) on the cases in (3.19) is again by chance. It just so happens that in all the three cases in (3.19), the base tone carries a high Register value [+upper], and that it is not difficult to lower this high Register value to a low one [-upper] so as to encode the lowered pitch value due to the falling intonation. Yet, when the underlying tone has a [-upper] feature specification, her analysis fails to work.

A similar problem arises if Yip's analysis is applied on the rising intonation on Clause 1. But first, let us return to Chao's original description of the data:

> "[In the sentence *nǐ xìng Wáng, wǒ xìng Lù*], the rising intonation in the further referring clause will make the rising 2nd tone *Wang* rise higher than usual and the falling intonation on the falling 4th tone *Lu* to fall lower than usual...But in *wǒ xìng Lù, nǐ xìng Wáng,* Lu will be pitched higher as a whole and still with a 4th tone contour, and *Wang* will be pitched lower as a whole, but still with a 2nd tone contour." (p. 39-40).

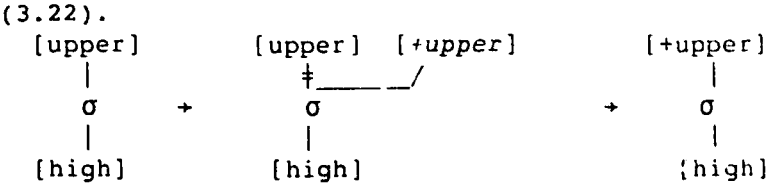
As have been shown earlier, two factors are involved here. One is intonation, which now rises at the end of the first clause (and falls at the end of the second clause). The other is tone, and there are a case of a rising tone (*Wang*) and a case of a falling tone (*Lu*). These two factors interact to yield four situations, as illustrated below:

(3.21).

	intonation	
tone	rising (first clause)	falling (second clause)
rising (T2)	1	3
falling(T4)	2	4

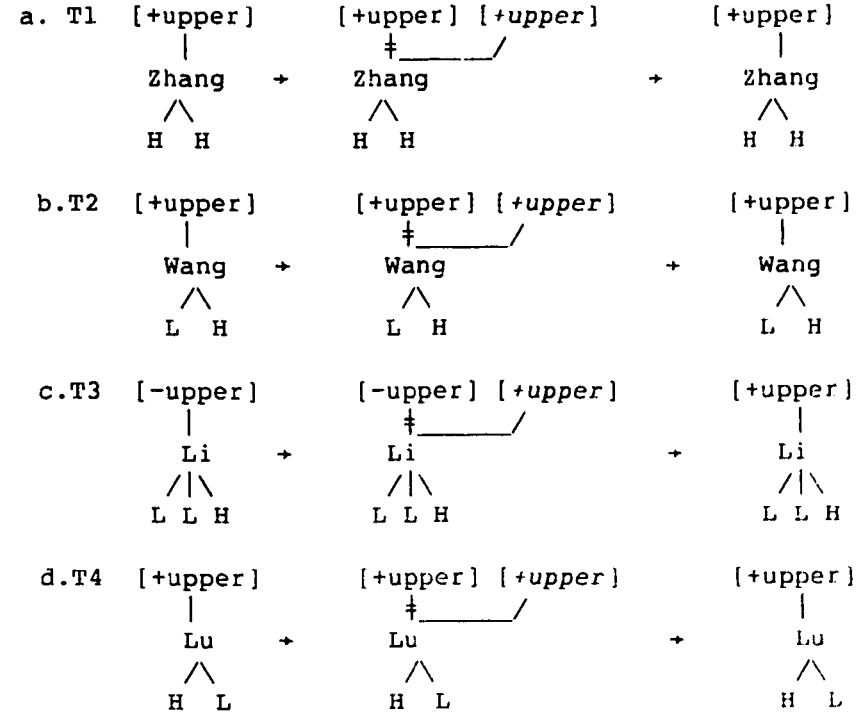
As was shown ealier, Yip's analysis for the falling intonation involves cases in 3 and 4 in the above. Now let us see how Yip's analysis would apply to the cases of the rising intonation (i.e. 1 and 2). Logically speaking, if Yip uses [-upper] to mark a low floating intonation in the case of a falling intonation, she would have to use [+upper] for a high floating intonation in the case of a rising intonation. Yip's analysis would then have a rising intonation rule as follows, which is an exact counterpart to her rule for the falling intonation in (3.18).

(3.22).



The examples shown below are the derivations via this rule on four Mandarin morphemes (the same four morphemes as seen in the previous case of the falling intonation), each carrying a different tone:

(3.23). URs SRs



Different from the falling tone cases where just one of the four tones poses a problem for Yip's analysis, there exist three tones here that pose problems for her analysis. This is manifested if the URs and the SRs in (a), (b) and (d) in the above are compared. In all the three cases, no change has occurred after the derivations. The question is again: what happened to the intonation - the rising intonation this time? It seems that just as in the case of the falling intonation on Tone 3, the goal to represent and thereby account for the pitch level raised by the rising intonation on the relevant tones is not achievable by this analysis.

The crucial problem here is probably attributable to a confusion of tonal analysis with analysis of intonation. With the same TONAL feature for both, the pitch imposed on the [+upper] tones by the rising intonation becomes simply out of the pitch range covered by Yip's tonal features.

What I have tried to show in this section is that the floating Register Feature used by Yip to encode intonation fails to adequately represent the intonation across all cases, and its legitimacy as a floating autosegmental feature thus becomes questionable.

3.4 Mandarin Case 4: "Intonation Particles"

Still another piece of Mandarin evidence provided by Yip for her Register Feature analysis suffers from problems similar to those seen in the previous Mandarin cases. As in the previous cases, her evidence here is again from Chao (1968). In his section addressing syntactic particles, Chao describes two intonation particles. He names these two particles "rising ending" and "falling ending" respectively. According to Chao, "these two intonational endings of phrases and sentences are of a very special morphophonemic nature (p. 812)." He then explains what this special nature is:

> "I used to treat these [endings] as part of Chinese sentence intonation, but later found it better to treat them as particles, since they do not affect the intonation pattern of the whole construction, but only the voiced part of the last syllable...[They] reside parasitically on the last morpheme by prolonging it for the length of a neutral-tone syllable on which to put a rising or falling ending." (p. 812)

A more formal description of the rising endings is given below in the column with the SRs:

(3.24).

	URs		SRs
a. 1st tone	55	→	556
b. 2nd tone	35	→	356
c. 3rd tone	214	→	2145
d. 4th tone	51	→	513

The following shows the shift from the meaning of the sentence without the rising ending (ii) to the meaning of the same sentence with the ending (i), the ending being signalled by the punctuation marks "!?".

(3.25). a. 1st tone on the final syllable
(i) qu lundun!?
"Do you mean that you are going to London?"
(ii) qu lundun.
"go to London."

b. 2nd tone on the final syllable
(i) bu xing!?
"Do you mean that it is not acceptable?"
(ii) bu xing.
"It won't do."

c. 3rd tone on the final syllable
(i) jiu yao zou!?
"Do you mean that you are leaving now?"
(ii) jiu yao zou.
"will leave soon."

d. 4th tone on the final syllable
(i) ni bu yao!?
"Do you mean that you don't want it?"
(ii) ni bu yao.
"You don't want it."

As indicated above, the rising ending expresses, among other things, the meaning of incredulity. In addition, there are two other characteristics about these intonation endings made obvious in Chao's description of them given earlier:

(3.26). a. The phrasal-final syllable affected is PROLONGED in duration.

b. This prolonged extra part, which constitutes a particle, is RAISED in pitch (thus, the name "rising ending").

Now let us see Yip's analysis for the intonation endings:

(3.27).

	URs				SRs
a.T1	[+upper]	→	[+upper]	→	[+upper]
	H H		H H *H*		H H H
b.T2	[+upper]	→	[+upper]	→	[+upper]
	L H		L H *H*		L H H
c.T3	[-upper]	→	[-upper]	→	[-upper]
	L L H		L L H *H*		L L H H
d.T4	[+upper]	→	[+upper]	→	[+upper]
	H L		H L *H*		H L H

Yip claims that the rising ending constitutes a floating H toneme only, its Register tier being unspecified. This floating toneme then gets associated to the Register Feature of the last tone in the phrase. There does not seem to be anything inappropriate regarding these technical procedures. Nor is there any obvious problem with regards to the encoding of the prolongation of the syllable in question. The addition of a high toneme does give the reading of a prolonged duration on the syllable (although Yip never explicitly formalizes her tonemes as timing units), and thereby achieves the result described in (3.26a). However, when one evaluates the results against (3.26b), one sees a problem. The problem is: where is the rise? In all three cases from (a) to (c), the rise is not encoded.

One may perhaps argue that the pitch of the rising ending (6 on Chao's scale) is after all not terribly different from its preceding toneme (5 on Chao's scale), and therefore may not be crucial at the phonological level. However, the fact is that just because of this slight rise as well as the prolongation, the meaning changes from one of a statement (3.25(i)s) to one of an interrogative sentence showing disbelief (3.25(ii)s). Thus, trivial as the difference is in pitch, it is, nonetheless, a meaningful one and therefore should be treated phonologically.

Here, one may ask the question of whether the prolongation alone would be enough to encode the rising ending. The answer to this question is no. In fact, results from some studies indicate that pitch is the most important factor in the categorial perception of tones. In certain instrumental studies (to be discussed below), for instance, Mandarin listeners have been found not to be sensitive to features such as duration in distinguishing tones across categories; the cue these speakers do make crucial use of is pitch variation.

One such study is found in Howie (1976) In one of his several experimental studies on Mandarin phonetics, Howie attempts to determine what the most essential characteristic a tone must have for it to be successfully recognized. To do so, he synthesizes Mandarin utterances with certain instruments and asks his subjects, native Mandarin speakers, to make judgements concerning the tonal categories of each syllable uttered. His result shows that pitch is the most crucial factor for correct judgements by these Mandarin-speakers. Other factors such as duration, intensity and vowel quality are not found to facilitate such perception:

> "The Mandarin listeners could apparently make little use of any features other than pitch as cues for the perception of tonal distinction." (Howie 1976, p. 245)

Howie also finds that regarding the other factors such as duration, intensity, vowel quality,

> "while they were audible when variation of the fundamental [frequency] was suppressed, they were not able to serve as supplementary cues for identifying the four tones in the absence of the pitch feature." (p. 242).

Howie's findings confirm the conclusions reached in an earlier study by Abramson (1962), who finds that in Thai, a language with a Chinese-like tonal system, pitch overrides the concomitant phonetic features such as duration, intensity and vowel quality in tonal perception.

If the results from these studies are valid, there seems to be another reason to reject Yip's analysis. These studies seem to indicate that if a Mandarin speaker perceives a rising ending, its perception cannot have been triggered by the prolonged duration alone; the raised pitch has to be present to make the difference. Yip's analysis, which does not represent the changed pitch but only the prolongation, hence fails to adequately represent the rising ending.

Besides the failure to represent the raised pitch, there seems to another problem in Yip's analysis. This problem lies in an apparent inconsistency in her analysis. As one may have noticed, Yip's analysis yields comparable results for Tone 1, 2, and 3 (3.27a-c). Notice that in all of these three cases, the analysis simply encodes the increased duration of the syllable without showing the rise in the pitch. In (3.27d) (i.e. the case with a fourth tone), however, the rise IS encoded. Not only is the rise encoded, it is encoded as higher than would be expected otherwise: Instead of a median pitch 3 (i.e. the end point of 513) -- the value described in the original data in (3.24), one sees, under her analysis, a value that is equivalent to the top pitch 5 (represented as a H tone in Yip) measured on Chao's scale.

By now, it should be clear that Yip's Register Feature system has not proven adequate in accounting for the intonation data in question; thus, these data cannot be used as evidence for the postulation of the autosegmental Register Feature (or the Tonal Feature for that matter).

Besides Mandarin-specific evidence as seen in the above sections, Yip also provided three general arguments for the positing of her register feature. In the remainder of this chapter, these arguments will be discussed.

3.5 *Yin and Yang Tonal Split*

One of Yip's important arguments for her register feature system draws on a well-attested theory of historical development of the Chinese tones.[9] Among the few things that are known about the history of the Chinese tones is that there are categorically four tones (or *shēng*) in Middle Chinese (around A.D. 600). This is clearly shown in the dictionary of Chinese characters, named *Qiè yùn*, compiled in A.D. 601 by *Fǎyán Lù*. This dictionary, whose entries are arranged by tonal categories,[10] is the chief source for the study of Middle Chinese, the language from which all the modern Chinese languages are supposed to have descended. The four Middle Chinese tonal categories as listed in this dictionary are given below:[11]

(3.28). Four Middle Chinese Tones

		Literal translations
a.	píng	(level)
b.	shǎng	(rising)
c.	qù	(departing)
d.	rù	(entering)

Sometime around or after the seventh century, these four Middle Chinese tones underwent a split into two tonal groups, traditionally referred to by Chinese scholars as the *yīn* and *yáng* groups respectively. Historically, syllables with *yīn* tones correspond roughly to Middle Chinese syllables with voiceless initials, while

syllables with *yáng* tones[12] to those with voiced initials. The well-known theory is thus that the split is, in the beginning, conditioned by the difference in the voicing of the initial consonants. As this voicing contrast was later lost in most of the Chinese languages (one exception is Wu), the phonetically conditioned tonal difference became phonemic. The result of the split is an enriched system of eight tones as shown below:

(3.29).

		yin	yang
a.	píng	yin-ping	yang-ping
b.	shǎng	yin-shang	yang-shang
c.	qù	yin-qu	yang-qu
d.	rù	yin-ru	yang-ru

After the split, history witnessed merging of these eight tones in various fashions in the emerging[13] vernaculars which were to become the modern Chinese languages as they are known today. Thus, the various tonal systems seen today in the various Chinese languages are largely results of these merging processes (cf. Wang 1987 and Wang and Cheng 1987).

Yip argues that her analysis with the Register Feature [upper] provides a better analysis for the historical facts. If this argument is tenable, and I suppose it is,[14] one may perhaps ask what the success on the part of this analysis in accounting for the historical development has to say about the adequacy of the same analysis in accounting for synchronic facts? Although it seems evident on the surface that the diachronic success is a good indication for the synchronic adequacy, the answer to the above question is not immediately clear to me.

I may perhaps approach the problem from another perspective: it is probably reasonable to assume that if the *yin* and *yang* register distinction which has been attested for Middle Chinese still largely exists in today's Chinese languages, one would have reason enough to conclude that the Register-Feature system of analysis which works well for the enriched Middle Chinese tonal system should also

work for those of the modern Chinese languages. Unfortunately, however, this supposition cannot be verified.

In fact, among the dialects of the Chinese languages, only a few seem to have kept the register distinction virtually intact. One example is the Yue dialect of Guangzhou (Cantonese). The vast majority of the dialects, however have lost this distinction to a greater or lesser extent. The dialects that show the greatest paucity of this distinction are the Mandarin dialects. Of the four splits of the four Middle Chinese tones, only one - the split of the *ping* tone - has survived in Mandarin, if the Mandarin dialect of Taiyuan is ignored, in which even this split has become obscure (cf. Wang and Cheng 1987).

Categorically, the four tones in the Mandarin dialects correspond, roughly speaking, to the *ying-ping*, *yang-ping*, *shang*, and *qu* tones of the Middle Chinese system. What is remarkable about this is that the *ru* tones, both *yin* and *yang*, no longer exist in the Mandarin dialects, any more than the *yin* and *yang* distinction in the *shang* and *qu* tones.[15]

Furthermore, it should be pointed out that the survival of the *yin* and *yang* distinction may be merely categorial. This means that the higher versus lower PITCH distinction characterizing the Middle Chinese *yin* and *yang* tones respectively may be lost totally, even though categorically they can still be classified into *yin* and *yang* groups. Just as Wang and Cheng (1987) observed in talking about the tonal development in question:

> "The values of the [Chinese] tones have changed a great deal over these 15 centuries - in different ways according to the dialect. Even though two dialects may have preserved intact a historical CATEGORY of tone, its modern VALUE may be quite different; e.g. it may be rising in one dialect while falling in the other dialect."
> (p. 227)

Following from this observation is another one: what was higher register tone in Middle Chinese may turn out to be lower register tone in modern dialect and vice versa; or what was higher or lower register tone may turn out to measure on the middle range. In Mandarin (i.e. MSD), for instance, the two *ping* tones, *yin-ping* and *yang-ping*, are no longer *yin* and *yang* distinguished in the sense that one is of higher register and the other lower register. In fact, they both belong to the higher register category, valued at (55) and (35) each. The same is true for Wendeng, another Mandarin dialect spoken on the east coast of the Shandong peninsula. In this dialect, both the *yin-ping* and *yang-ping* tones, (51) and (55), belong to the *yin*, or higher register category. Xi'an, a western Mandarin dialect, has the opposite situation. Both of the *ping* tones, *yin* and *yang*, fall into the lower register category, with the values of (31) and (24) respectively.

Jinan Mandarin, spoken in the capital city of Shandong province, presents evidence of another sort: the *yin-ping* tone with a value of (213) is no longer *yin*, or higher-registered, while the *yang-ping* tone, with a value of (51), is no longer *yang* or lower-registered. The same is true for yet another Mandarin dialect, that of of Lingbao, whose *yin-ping* tone, (31), is no longer higher-registered and whose *yang-ping* tone (35) is no longer lower registered. Although I have only listed tones from the Mandarin dialects, the same is true in the tonal values of a vast majority of the other dialects of the Chinese languages.

However, the worst difficulty for the Register-Feature tonal system is not that one ancient tone of one register has become a tone of another register, but that many of the ancient tones have acquired a value that either occurs right at the middle point (3) or goes across it, covering both the upper and the lower registers.[16] As it was shown earlier, one of the Mandarin (i.e. Standard Mandarin) tones has the value of (51), covering the whole range of both registers. Shanghai, a dialect of Wu, has two such tones (42 and 24) among its five tones. Other dialects with such tones include the Wu dialect of Yongkang with two such tones (52 and 24) among its six tones, Wenzhou with two such tones (24 and 42) among its eight tones, and Suzhou with four such tones (24, 41, 513 and 3) among its six tones; the Min dialects of Xiamen with three such tones (24, 51 and 33) among its seven tones, and Fuzhou with three such tones (41, 342 and 24) among its seven tones; and the Yue dialects of Tengxian with three such tones (42, 24 and 33) among its seven tonal values and even Guangzhou (Cantonese) with two such tones (24 and 33) among its seven tonal values.[17] According to Packard's (1989) calculation, 33% of all contour tones in the Chinese dialects go across the middle range.

Largely because of its inability to represent these tones and to capture generalizations about certain tonal processes, Yip's Register-Feature system defining four levels of tones has very recently been explicitly (e.g. Chan 1991 in her analysis of Danyang and Zee 1991 in his analysis of Shanghai) or implicitly (e.g. Shen 1985, Jin 1986 and Selkirk and Shen 1990 in their analyses of the Shanghai tones) rejected by phonologists of Chinese. What is adopted instead is a system that distinguishes three levels (i.e. L M and H).

What the above discussion seems to suggest is that due to the large-scale merging processes, especially to the shifts of the tonal values in various directions, the historical register, or *yin* and *yang*, distinction is largely obliterated in

the modern Chinese languages. This seems to be particularly true for the Mandarin dialects in which the distinction can be said to have become completely opaque.

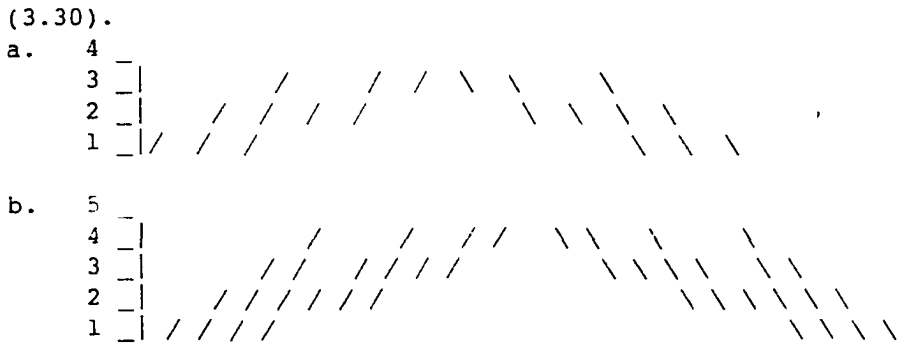
What then can be concluded at this point? Probably that the existence of the historical register distinction provides no evidence for its existence in the modern dialects, and the fact that an analysis that is successful in accounting for the historical facts does not necessarily provide *a priori* reason for its use in the analysis of the synchronic data.

This conclusion seems to be in line with the argument that historical evidence is irrelevant in the determination of a contemporary native speaker's grammar. Though it may be used to trace the origin of certain features already established in today's grammar, it cannot provide direct motivation for the establishment of such a feature in the contemporary grammar. Just as it is put in SPE, "[Historical explanation] is obviously irrelevant as far as the linguistic competence of the native speaker is concerned (p.373)."

3.6 The Constraining of the Generative Power

Besides historical evidence, Yip also supports her recognition of the Register Feature by what I may call a "power constraining" argument. In particular, Yip criticizes other systems as being too powerful. A previous analysis that distinguishes four level tones, argues Yip, predicts six rising and six falling tones, whereas another analysis that distinguishes five level tones predicts ten rising and ten falling tones. Yip illustrates this problem as follows:

(3.30).



What is illustrated above is (a) the set of six rising tones: {(12), (13), (14), (23), (24), (34)} and six falling tones: {(43), (42), (32), (41), (31), (21)} possible in a traditional four-level system, and (b) the set of ten rising tones: {(12), (13), (14), (15) (23), (24), (25), (34), (35), (45)} and ten falling tones: {(54), (53), (43), (52), (42), (32), (51), (41), (31), (21)} possible in a traditional five-level system. Yip contends that no language has so many underlying contrasts of contour tones, but a language usually contrasts no more than two rising and two falling tones. Her analysis is thus superior to those previous ones since it is confined exactly to the generation of that limited number of four contour tones.

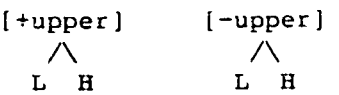
Recall that the curb of the generative power of her analysis is achieved with the postulation of the following constraint on the distribution of her Register Feature (see also § 2.7).

(3.31) "[The] register [feature] remains constant over the morpheme." (p. 126)

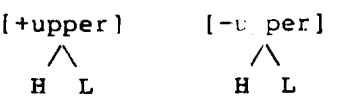
One thing about Yip's success in the constraining of the generative power of her analysis that should be noted here is that the constrained power does not come from the internal mechanism of her feature system, but from an external statement as given above. Nevertheless, with this constraint, Yip's system is indeed successfully constrained with but two technical problems: one, the constraint should really be stated more precisely as being over the tone bearing unit (or over

the syllable which is the tone bearing unit in the Chinese case)[18] rather than the morpheme, since the morpheme may contain more than one syllable, and is not exactly the unit that always bears one single tone. Two, as noted by Bao (1990), Yip's analysis needs another constraint for her goal of limiting the number of contour tones to four to be achieved. This extra constraint would prevent the tones from getting unreasonably long, as nothing in Yip would prevent a tone such as HLHLHL, or even longer to occur. Considering these two as merely minor problems, let us see the four contour tones that are allowed in Yip:

(3.32). a. rising tones



b. falling tones



Here, a question might be asked: has the constrained power been achieved at a sacrifice? Packard's (1989) study suggests a positive answer to the question.

3.6.1 Packard's (1989) Study

In a more formal way, Yip's constraint as stated in (3.31) confines contour tones in any given language to appear exclusively either in the higher [+upper] range or in the lower [-upper] range. This means the exclusion from the grammar of a representation such as the following:

(3.33).

In the above, the register feature [upper] alternates its value within the domain of the tone-bearing unit, in violation of Yip's constraint as stated in (3.31), and is thus ruled out from the grammar.

In an attempt to analyse tones of some of the Chinese dialects, Packard finds this analysis by Yip to be over-constrained. Based on a survey of the Chinese dialects described in *Hànyǔ Fāngyán Gàiyào* (An Introduction to Chinese Dialects) by Yuan (1960), Packard finds that 33% of all the contour tones bestride the mid-voice range, covering thus both the [+upper] and [-upper] ranges. He argues that although many of these contour tones may be considered derived, "it does not seem reasonable to assume that all of them are derived (p.20)." He demonstrates this point with a tone from a Wu dialect called Suzhou.

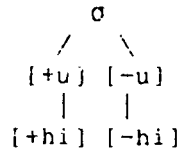
In Suzhou, Tone 3^{19} is a tone that is (52) in value. According to Packard, this tone "never changes shape in word- or phrase-initial position ...In non-initial position, Tone 3 undergoes largely the same tonal alternations that all the other tones ... undergo in that position (p. 20)." He therefore concludes that "there seems little reason to assume that Tone 3 is derived (p. 20)."

Packard then conducts analyses for a number of these Chinese dialects including Mandarin, Wuxi (another Wu dialect), Xiamen (a Min dialect), and Meixian (a Hakka, i.e. Kejia, dialect), and shows that data from these dialects can only be handled properly if Yip's constraint as stated in (3.31) is removed, and the analysis in (3.33) be accepted and incorporated. Specifically, he believes that

> "all tones (with some exceptions, such as for 'checked' or 'neutral' tone) are given two values for register in underlying form... The two identical registers are collapsed by a universal rule called 'register collapse'... Tones which do not undergo register collapse remain double-specified for register and possess a BRANCHING register structure." (p. 21)

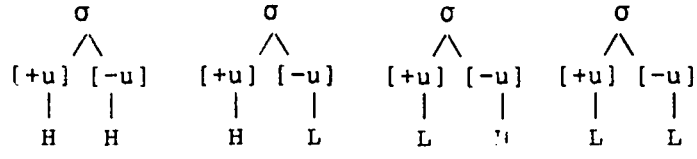
By "branching register structure", he means exactly the structure as shown in (3.33) which is ruled out by Yip's constraint. An example of Packard's representation of the Chinese tones may be taken from his representation (p.22) of the fourth Mandarin tone (51), which is given below:

(3.34).

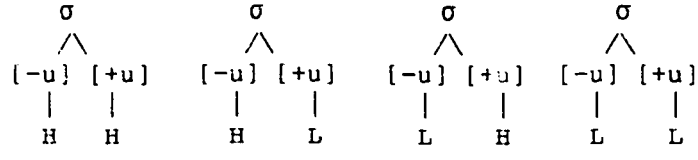


Interestingly, Packard simply extends Yip's analysis in an effort to solve the problems he observes in it, without being aware that his extension has in fact undermined the very cornerstone of Yip's analysis. Yip's analysis, when thus extended, has in fact lost its advantage of being more constrained than other analyses, as it is no longer confined to the prediction of four and only four contour tones. Thus extended, Yip's analysis also predicts six rising and six falling tones, just as a traditional four-level analysis does as illustrated in (3.30b). The following illustrates the extra contour tones, besides those four described earlier in (3.32), which would be generated by Yip's analysis, should it be extended according to Packard.

(3.35). a. falling tones



b. rising tones



What I have demonstrated so far in this section is that Yip's positing of the Register Feature has not made her analysis adequately or more easily constrained, since with the constraint placed on the register feature, her analysis becomes virtually "disabled".

3.6.2 The Function of a Feature System

Having found that Yip's analysis is constrained in such a way as to eliminate the descriptions of actually occurring contour tones, I may perhaps turn to the very basis of Yip's "power constraining" argument and deal with a more fundamental question: Why is it necessary to confine the predictive power categorically to exactly two rising and two falling tones? The answer to this question has to do with another argument Yip advances in support of her register feature: that virtually no language contrasts more than two rising or falling contour tones. Supposing Yip's observation is true, one may then want to re-phrase the question and ask instead: does one want to limit the predictive power accordingly (i.e. to the prediction of two rising and two falling tones only)? The answer does not seem to be easy and transparent. In fact, one needs to go to the very base of generative phonology and inquire into the fundamental function of a system of features.

As is well-known, a system of phonetic features should be able to define possible speech sounds in languages. It should, as SPE put it, "represent phonetic capabilities of man (p. 299)." A system of features that describe vowels, for example, should describe all the vowels that occur cross-linguistically. A look at a usual IPA chart reveals that there are something like a dozen or more simple vowels found in languages in general. A system of features should be able to describe and distinguish all of these vowels.

In phonological theory and practice, the fact that no language contrasts as many vowels as are listed in the universal vowel inventory has not been used as

evidence to limit the universal feature system to the definition of just the maximum number of vowels any given language can possibly contrast.

Let me spell this argument out more concretely. Suppose that there were twelve members in the set of vowels found in all languages, and meanwhile, the maximum number of vowels contrasted in any given language were a subset of the twelve - seven, for instance. The question is: should the phonetic descriptive apparatus be thus controlled so that only seven vowels, not twelve, can be contrasted? It is clear that the answer is no. That a language can contrast maximally seven vowels provides no base for the choice of a feature system which can only define categorically seven vowels. An adequate system of features should, rather, be able to distinguish twelve rather than seven vowels.

Similarly, the generative system of features that describes consonants is in fact able to define all the consonants attested in human languages, not just the maximal number any given language can possibly contrast. It seems expected that the overall set of consonants in human language is much larger than the set found in any given language, and that no language actually contrasts the same number of consonants as contrasted in this overall inventory of consonants found across languages. Conversely, the fact that no language distinguishes the same number of consonants as found in the overall inventory is no argument for limiting the feature system to the definition of just the number of consonants found maximally in any one language.

The same argument for the vowels and consonants is evidently applicable to a system of tonal features. It may be very well true that hardly any tonal language contains more than four contour tones, but this fact should provide no argument against the configuration of a system of tonal features in such a way as to define all the possible contour tones in languages. Ideally, a theory of tonal features

should be able to distinguish various tones in all tonal languages. If not for any other grand purposes, a theory of such a feature system at least provides a necessary meta-language with which cross-language comparison of contour tones can be made.

It should also be pointed out that by confining her model to the distinction of four tonal categories, Yip in effect confined her model to the definition of just four TYPES of tones altogether. It would seem quite unlikely that the overall number of contour tones in human language should be as small as four, for although it might be true that no language contrasts more than four contour tones, the four-way contrast does not necessarily have to always involve the same four tones; in other words, a system that defines more than four contour tones does not necessarily cause more than four tones to be contrasted in any given language.

It seems that Yip's fundamental problem lies in that she confuses phonetic features with phonological features. In particular, she tries to develop what amounts to a UNIVERSAL set of PHONOLOGICAL features for tonal analysis. Such a set of features clearly does not exist. Some basic tenets of theoretical phonology suggest that a certain subset of the universal set of phonetic features may turn out to be phonological, contrastive features in certain language while not others, and that all phonetic features are not phonological features in all languages. An adequate feature system for tone should contain a set of phonetic features that can define all tones in languages the system is supposed to cover.

To conclude, Yip's system of tonal features, which is capable of defining just four types of contour tones, suffers from under-generation, and therefore her power-constraining argument for the adoption of the register feature is undermotivated.

3.7 *Physiological Considerations*

It is rightfully pointed out by Yip that "any system of distinctive features must of course be grounded in an understanding of their articulatory and acoustic correlates (p.155)." According to her, attempts have been made in this direction in devising a tonal feature system which reflects the correlation between pitch and certain articulators (Halle and Stevens 1971, and Brown 1965). While making no claim that her analysis is better than these attempts in this regard, she does point out that none of the attempts is successful. She then concludes that "phonetic motivation for any system of tone features must await a clearer understanding of the mechanism of production (and perception) (p.157)." I think Yip is correct here.

However, I still have some doubt which I wish to raise here regarding the physiological correlates of Yip's tonal system. It seems that a commonly accepted theory is that tone is produced by the laryngeal muscles (cf. Ohala 1978), the vocal cords in particular. But, whatever the relevant muscles, and wherever they are located in the vocal tract, it seems reasonable to assume that Yip's two tonal features including the Register Feature are related to the physical movement of the same set of muscles (or whatever), and therefore share the same articulatory correlates, given the fact that they both are tonal features covering the same pitch range (that is, 1-5 on Chao's scale).

A problem seems to arise here: if Yip's Register Feature [upper] is conceived as having an independent life from the Tonal Feature [high], it is not clear how the same set of muscles can in fact effect two simultaneous but distinctive kinds of TONAL behavior. Without further evidence from articulatory or perceptual studies, I have to remain skeptical about the physiological reality of the Register Feature [upper], assuming that the Tonal Feature [high] is justified. As I am no expert at all on the subject matter under discussion, I do not intend to pursue the matter

further in this study. I will content myself with the mere raising of the question, in the hope that further research may shed light on it.

3.8 Notes to Chapter 3

1 Unfortunately, however, I got to know Zee's position only through the abstract of the paper Zee presented at a 1991 Chinese linguistics conference. My recent effort to trace the author and his paper has failed. I therefore have no details about Zee's analysis.

2 This is my nearly verbatim translation of Xu. The original remark is in Chinese, which is provided here in *pinyin*.

> "qīngshēng yě jiào qīngyīn, tā de tèdiǎn shì fāyīn qīng ér duǎn." (p. 220)

3 See Chen (1984) for a different view.

4 The names or the exact syntactic functions of these particles may be controversial. However, the controversy should not affect the present discussion.

5 Diachronically, though, they have been found to be derived from full-toned morphemes (Cheng, p.65).

6 She does not mention the source of her data here. Presumably they are from Qi (1956), since they identify with Qi's. The data are almost identical to Chao's as well, which are given below (Chao 1968, p.36):

Tone 0

Tone 1	(55)	2
Tone 2	(35)	3
Tone 3	(213)	4
Tone 4	(51)	1

The only difference between Chao and Yip's data shows up on the neutral tone after Tone 1. In Chao it is measured at 2 while in Yip's citation, it appears as 3. I believe the difference is trivial and therefore negligible.

Notice also that in (3.3), Tone 3 is represented with a value of (21) rather than (213) (See § 2.7 for her analysis of this tone). The examples shown there are mine.

7 For a more detailed analysis and more arguments of the neutral tone, please see § 6.7.

8 I would not be surprised if there are researchers who do not agree in treating them all as identical neutral toned syllables. In fact, my intuitive feeling says that these post-verb syllables (except *le*) still carry distinct though reduced tones, the reduction being due to lack of stress, and no matter how much tone modification occurs in each case, these syllables are not produced with an exactly identical contour or pitch. However, this perception may be purely phonetic in nature. Namely, phonologically, these tones may very well be considered to be the same neutral tone.

9 The same argument is made in Bao (1990) as well.

10 According to Wang and Cheng (1987), the earliest discussion of the four Middle Chinese tones is due, however, to Yue Shen (A.D.441-531).

11 Except that all the syllables that ended in a stop (i.e. /p, t, k/) were grouped together under the *ru*-tone category, no one can tell with certainty what the names of the tonal categories meant when they were adopted. This means that nothing is known about the values (or shapes) of these tones.

12 Hereafter, I may omit the tone diacritics on the Chinese words that have already been mentioned to avoid unnecessary tedious post-script setting involved in the word processing of the present dissertation.

13 Splits also happened in some dialects (notably the Cantonese dialect of Yue) during the tonal development, but to so small an extent that they are negligible in the present discussion (also see Wang and Cheng 1987).

14 Due to the scope of her historical analysis, I will not go into it here, but sim-

ply assume that it is adequate.

15 The actual development of the four Mandarin tones involved a complicated series of processes. The loss of the *ru* tones (both *yin* and *yang*), for example, was due to the loss of the consonantal endings of the *ru*-toned syllables. These syllables consequently merged with their voicing counterparts of the *ping* tones. For a more detail discussion of the merging processes, the reader is refered to Wang and Cheng (1987).

16 This issue is further addressed later in § 3.6.

17 All the tonal values given above are from Norman (1988), Chapters 8 and 9, except those for Jinan and Wendeng, which are from my own auditory impressions.

18 Chinese languages are divided into syllable tone language and word tone language. Shih (1986) makes it quite convincing that Mandarin is a syllable-tone language in which the tone bearing unit is the syllable rather than the morpheme, in spite of the Mandarin morpheme being largely co-extensive with the syllable.

19 The use of a number for tone reference is the same here as in Mandarin.

Chapter IV

REPRESENTATION OF MANDARIN TONES

One argument that has enjoyed unanimous acceptance among current researchers of Chinese phonology is that Mandarin tones should be represented autosegmentally, but exactly what this autosegmental representation should look like is a question that has not received much attention in current phonological studies. As is revealed by the earlier review of literature, virtually the only non-linear analysis available is the analysis of Chinese tones developed in Yip (1980a), and later modified by the author herself (1989a,b). However, Yip's framework, as has been demonstrated in the last chapter, has not been found motivated for Mandarin tones. The purpose of this chapter is thus to offer an alternative non-linear representation of these Mandarin tones.

4.1 Preliminaries

By now, it should be a familiar piece of information that Mandarin has four phonemic tones. Among them, the first is high and level (55, on Chao's scale), the second rising (35), the third low and dipping (213), and finally, the fourth falling (51). These four tones will hereafter be referred to as Tones 1, 2, 3, and 4 respectively, and the present chapter will address primarily the representation of these four phonemic tones.

It should be noted that the values for the four tones as given above are their values in citation. In certain contexts, these tones may appear in sandhi forms different from their forms in citation. Traditionally, these sandhi forms (or sandhi

tones) are seen as derived from the citation forms (or citation tones) through tone sandhi processes. Basically, four sandhi processes have been identified in the literature. This chapter will focus primarily on the representation of the four phonemic tones; Chapter 6 will be devoted to a detailed discussion of these tone sandhi processes.

As I have already shown (especially in § 3.1), Mandarin has another tone commonly referred to as the neutral tone, in addition to the four phonemic tones. Different from the other four tones, the neutral tone is still generating much debate as to whether it is underlying or derived, or how it acquires its surface tonal values. This neutral tone, hereafter referred to as Tone 0, will be further addressed in a later chapter.

4.1.1 A Working Definition for the Term "Tone"

In order to delineate a theory of tone, a few key terms need to be defined at the outset. One of them is the term *tone*. Much confusion exists in the literature concerning the denotation of this term. As will be demonstrated below, this confusion is attributable to the fact that a tone is not always atomic. Indeed, within the framework of non-linear phonology, a tone may be decomposable; it may have an internal structure, consisting of a sequence of "tones" itself.

To confound the picture even more, the domain within which a tone may occur varies from case to case. As once noted by Leben (1978, p.177), the domain of the tone varies among the units of the phonological word, the morpheme, the syllable, the mora, and the segment. Recently, the set has been extended to include the unit of the lexical word (Shih 1986, Zhang 1989). In autosegmental terms, such a domain is called a tone-bearing unit (TBU).

The confusion stemming from the TBU (domain) variation can be demonstrated by two of the scenarios of the TBUs. In one of the scenarios, the TBU is the

syllable. In the other, the TBU is the lexical word. Incidentally, these are the two scenarios that have been found to exist in the Chinese languages. Based on Shih's (1986) distinction between word-tone languages and syllable-tone languages, Zhang (1989), for instance, finds that the TBU in Mandarin is the syllable while that in Shanghai is the word.

In a language like Mandarin, in which the TBU is the syllable, a contour tone such as Tone 4 may typically be represented within non-linear phonology as follows:[1]

(4.1).

In this representation, there are two levels involved. Confusion thus arises from the fact that the units at these different levels have both been referred to as "tones". Specifically, the term "tone" has been used alternatively to mean (1) the whole falling contour as a unit, and thus the expression "a falling *tone*," and (2) one of the components of the contour, and therefore, the expression "a sequence of H and L *tones*". This situation gets even worse when the TBU is the lexical word, for the word may contain more than one syllable, and each syllable, in turn, may bear a tone with an internal structure. An illustration is provided below based on a lexical word tone from Shanghai:

(4.2). A Shanghai Word Tone: 42+ 24

2-syllable	3-syllable	4-syllable
42-24	42-42-24	42-42-42-24

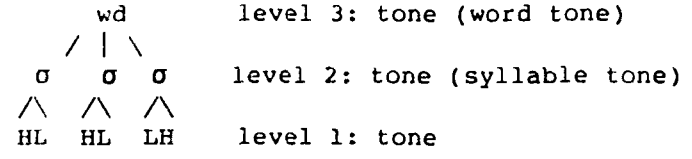
Here, χ+ = any number of χs from 1 up

As mentioned previously in § 2.8, a lexical word tone in Shanghai is usually either extended or compressed to fit the entire domain of a lexical word, no matter how many syllables the domain of the word contains. Within such a lexical

tone, three levels of tonal units are discernable. From top to bottom, one sees the units at the word level, at the syllable level and at the level of the components making up the unit of the syllable level.

What is confusing regarding these units is that all of them have been referred to as "tone". Taking the trisyllabic word from (4.2) as an example, I can display the problem as follows:

(4.3).



The case as illustrated in both (4.2) and (4.3) clearly suggests that it is wise to define the units at these three levels using separate terms if one wishes to avoid unnecessary confusion in the use of the term "tone" as well as any further confusions that may result from it.

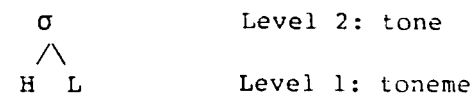
Hence, in the development of my analysis, I will explicitly adopt the terms "melody" and "toneme" to refer to units at the top and bottom levels respectively, while reserving the term "tone" for the exclusive reference to the tonal unit at the middle (or the syllable) level.[2] Namely, the following definitions are assumed:

(4.4). A working definition for the terms

"tone," "toneme" and "melody"

Under such a scheme of definitions, each of the four Mandarin phonemic tones as well as its neutral tone, all of which are directly linked to the domain of the syllable, will be referred to as "tones". Meanwhile, the component parts of each of these five tones will be referred to as "tonemes." What I mean is that (4.1) is labelled as follows:

(4.5). Mandarin tone and tonemes:

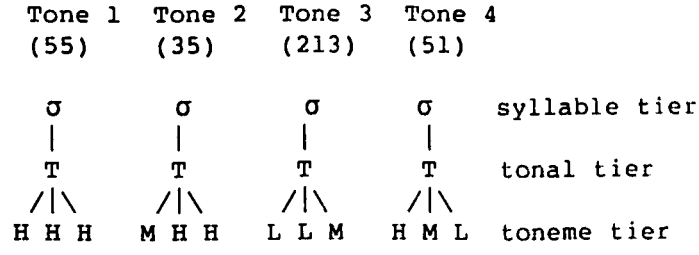


Having defined the term "tone" and other related terms, I am now ready to spell out my analysis of Mandarin tonal representation. I will begin with a brief overview of the analysis, followed by an exposition of its major features and their supporting arguments.

4.2 The Representation of Mandarin Tones

The analysis proposed here for the four Mandarin tones consists of essentially the following geometrical structures; these structures are considered the underlying representation of the four tones:

(4.6). Representation of the Four Mandarin Tones



Basically, there are three tiers involved in the representation, the syllable tier, the tonal tier, and the toneme tier. While the tonal tier is treated as being co-extensive with the syllable tier, the toneme tier is seen as linked directly to the tonal tier, and only indirectly to the syllable tier.

Regarding the treatment of the syllable and tonal tiers as being co-extensive with each other, one may argue that it would seem more precise to link the tonal tier with the voiced part of the syllable only (rather than with the whole syllable itself), given the fact that the tone is only observed over that part of the syllable. However, it can also be argued that in spite of the observed fact, it is highly likely

that the tone is linked directly to the syllable as a whole in the phonological grammar of the native speaker.

The indirect linking between the tonemes and the syllable via an intermediate T node is meant to accommodate the essentially correct observation described in Yip's two recent studies (1989a,b) that the Chinese tones, contour tones in particular, should be treated as units at one level of the phonological representation, in addition to being treated as being composed of level tonemes. While Yip's motivation for the T node comes mostly from Wu dialects (see § 2.8), I will later in this chapter provide a piece of evidence from Mandarin to further support Yip's treatment.

Generally speaking, there are three overall characteristics about our analysis, which stand out as distinct from previous analyses of Mandarin tones. While detailed exposition of them will be provided in the sections that follow, a brief introduction is offered here. The first of the three characteristics is that the present analysis assumes that THREE tonal levels, high, mid and low, rather than four (as in Yip 1980a) or five as in Chao (1968) are contrasted in Mandarin. In order to represent such a three-way contrast, two commonly used traditional features, [+/-high] and [+/-low], are adopted.[3] The two features combine to yield three tonal levels as desired:

(4.7).

	[high]	[low]
high	+	-
mid	-	-
low	-	+

These three tonal levels can be represented by three tonemes, H, M and L, as shown below together with their featural definitions:

(4.8).

level	tonemes	feature representation
high	H	[+high, -low]
mid	M	[-high, -low]
low	L	[-high, +low]

Second, the present analysis assumes that each of the four Mandarin tones is composed of a designated combination of three (rather than two) of these tonemes. Tone 1, being a high level tone, is represented in terms of a flat series of three H tonemes. Tone 2, being a mid rising tone, is represented as starting with a M toneme, continuing to a H toneme, and remaining high for the length of another toneme H. Tone 3, a low tone, is configured as starting from L continuing through another L, and then rising to M. And finally, Tone 4 starts at H and then drops through M to L.

The third characteristic is unique since it is a traditionally phonetic notion that has not been explored in the phonological analysis of tones. In particular, this notion is one of timing, comparable to the notion of timing used in characterizing Cs and Vs, - units on the skeletal tier in languages which have been typologically grouped together under the term "quantity-sensitive languages".[4] I claim that tonemes in Mandarin function as such timing units. Essentially, I argue that Mandarin, though quantity-insensitive at the skeletal level, is nevertheless quantity sensitive at the tonemic level. It should be noted, however, that with this claim, the traditional classification that Mandarin is a "quantity-insensitive" language needs to be modified, and the weight theory of the non-linear phonology is in need of review, for the parameter of quantity-sensitivity becomes now no longer exclusively relevant to the skeletal tier, as it has always been considered to be. This issue of the timing units will be addressed in more detail in the next chapter.

Apart from the three unique features just mentioned, my analysis adopts a few related assumptions made in previous analyses. These include, as briefly discussed earlier, that (1) Chinese tones should be treated as a sequence of level tonemes, following Woo; that (2) Chinese tones should be treated as atomic units as well as sequences of level tonemes, following Yip (1989a,b); and that (3) the

TBU in Mandarin is considered to be the syllable, following all of the previous studies except Woo. Besides these three assumptions, there is yet a fourth one - that (4) Mandarin citation (or utterance-final) tones are the underlying forms, following the traditional approach first proposed in Chao.

Among these four assumptions, (1) and (3) have enjoyed relative unanimous acceptance by phonologists of Chinese in general. Hence, I will not address them further in this study. Neither will I be concerned much with (2), as it is very cogently argued for in Yip's (1989a,b) two recent papers (see § 2.8), although I will provide further evidence from Mandarin to support Yip's treatment. Since (4) remains an issue of debate, I will provide explicit arguments in the following sections to justify the adoption of it.

4.2.1 The Nature of the Toneme

First, I would like to argue for the postulation of three tonemes (rather than two) in the underlying representation. To do so, it is first necessary to probe further into the nature (or the basic function) of a toneme. It seems that there are at least two parameters that can be used to characterize a toneme. The first one is, of course, the parameter of the pitch level, as a toneme always specifies a pitch level. A sequence of tonemes will tell the direction of the pitch to indicate if the tone containing them is a level or a contour one and, if contour, what direction the contour takes. For instance, a sequence of LH tonemes would tell that the tone is marked by a change in the fundamental frequency in the course of its utterance, and reveal that the tone is a contour rather than level tone. Or, on the other hand, a HH sequence would mean no change in the fundamental frequency in the tone, revealing therefore a level tone.

The second function of a toneme, however, is not as well understood. It concerns the fact that a toneme can potentially serve as a timing unit in languages

such as Chinese, just as the skeletal elements of Cs and Vs can in others. This function of the toneme, little known as it is, has in fact been implicitly assumed in studies of tones. For instance, there is a reason to believe that several previous studies, such as Chao, Woo and Yip, treat, largely unintentionally, the component parts (tonemes) of a tone as timing units.

The numerals in Chao, for instance, can be interpreted as indicating duration. In his system of tonal representation, Chao uses a "geminate" (55) rather than a single (5) to denote the high level Tone 1. Such a representation of Tone 1 evidently serves to indicate that a tonal sequence represented by geminate numbers is somehow longer than one represented by a single number. Otherwise, there is no point in using two identical numbers (55) rather than a single one (5) in this case. In other words, if pitch level is all there is to represent, a single (5) would work just as well as a dual (55). Treatment of tonemes as timing units is also found in Woo, Yip and Bao. In what follows, I will omit Bao in the discussion, since Bao is essentially the same as Yip as far as the system of tonal features and tonal levels are concerned.

As shown earlier in the literature review, Woo and Yip represent Tone 1 as HH rather than simply H (although Woo "stretches" the vowel length at the same time so that each of the tonemes can be carried by its own vowel, as required by the segmental framework she works in). Furthermore, besides Tone 1, Tone 3 is represented in these studies by two identical tonemes. Specifically, Tone 3 is represented in them as LL rather than simply L. Such a treatment of Tone 1 and Tone 3 suggests that tonemes have been used as timing units. In fact, the following remarks made by Yip further suggests this.

> "The last question to be discussed concerns the possibility of contrasts between single tones and sequences of identical tones...I see no reason to rule out the possibility of a contrast between, say HH and H underlyingly... I would suggest that there MAY be a distinction between LL and L that shows up in the timing of such sequences as LLH and LH on a single syllable." (p.33).

Unfortunately, these previous analyses, which endow, to a greater or lesser degree, the integral components of a tone with a time-denoting function, have only done so without seriously considering the consequences of such a treatment, and without being explicit about it. These consequences will be further discussed in the next section.

4.2.2 The Duration of Tone 3

If an analysis uses tonemes, implicitly or explicitly, as timing units, a logical question emerges: How many tonemes should be strung together to represent a given tone, or to represent the duration (rather than merely the pitch contour) of the tone? Again, let us consider the studies of Chao, Woo and Yip, which are relevant to the present topic.

(4.9).

	Tone 1	Tone 2	Tone 3	Tone 4	
a. Chao					
	55	35	213	51	UR: in final position
	55	35	21 35	53	SR: in non-final position
b. Woo and Yip					
	HH	LH	LL	HL	UR: in non-final position
	HH	LH	LLH	HL	SR: in final position

Assuming the timing function of the tonemes, two inferences contained in all these three studies can be drawn from the above illustration. One, all tones, except Tone 3, remain the same in duration no matter where they occur. Namely, they are always "two-tonemes long" wherever they appear. Two, the third tone is longer by about 1/3 in a final position than in a non-final position.

These two assumptions, seemingly plausible at first glance, lack empirical support. First of all, evidence from some instrumental studies shows that other things being equal, all Mandarin tones, not just Tone 3, are significantly longer in utterance-final position than elsewhere. Secondly, these studies also show that while Tone 3 may perhaps be the longest among the Mandarin tones in final position, the difference is not as great as previously assumed.

The observation that all Mandarin tones are longer in duration in an environment preceding an empty time space rather than in one preceding another tone has been made in several studies. Chao (1968), for instance, remarks that "a two-syllable compound or phrase [standing alone] will have a slightly greater stress on the second syllable...This is true of any combination of tones... (p.29)." Further, in describing the effect of a stress on a tone (which is co-extensive with the syllable that bears it), Chao maintains that "stress will enlarge ... the length of a tone (p. 29)." Chao's correlation between stress and duration is confirmed in a fairly recent study in Shen (1991), whose perceptual and acoustic study of Mandarin rhythm reveals that "the mean timing of the stressed syllables is 237.81 ms.; that of unstressed syllable is 145.14 ms.", and that "the stressed syllable is significantly longer than unstressed ones" (from her abstract).

Putting all these remarks together, Chao's theory becomes clear: that the syllable situated before a pause has a longer duration than a syllable that is not found in that particular context. This implies, among other things, that it is not just Tone 3 but every one of the Mandarin tones that is longer in final position.

More evidence is found in another study by Shen (1990b). In this study, Shen attempts to measure the low level co-articulation effect exerted on a tone by tones on its two sides. She records native speakers' pronunciations of a sequence of three identical syllables /pa pa pa/ in various tonal combinations. Before being recorded, the Mandarin-speaking subjects are specifically cautioned not to stress one syllable more than another in order to control stress-related modification of the tones to a minimum. Shen finds that in spite of the caution, the utterances made by these speakers are still "produced following the most frequent stress pattern of trisyllabic words or phrase, that is, with the last syllable most stressed and the second least stressed (p. 283)". If stress does indeed manifest itself as an enlargement of duration, Shen's finding confirms Chao's observation that the final syllable or tone is longer than one occurring in other contexts. Notice that Shen's finding suggests that all Mandarin tones including Tone 3 are longer in utterance-final position. Similar observations to those made in Shen are made in Yan and Lin (1988) as well.

While all the above studies suggest that a tone is indeed longer utterance-finally, one still need to answer one more important question to make the point under discussion: How significant is the increased duration? Is it an insignificant, low level duration modification, or is it significant enough that the phonological representation should reflect it? Findings from Shih's (1987) Bell Laboratory report may perhaps help us make the decision.

In Shih's report, it is shown that Tone 0 takes 30% more time in a final position than when in a non-final position. In addition, she reports that a vowel, measured to be 300 millisecond (ms) long when the syllable that contains it is in an utterance-final position, lasts only 150 ms (i.e. 50% less) when the syllable is in a non-final position. Given that a tone is largely co-extensive with the vowel rather

than with the preceding consonant of the syllable (Howie 1974), this finding should mean that a tone is much shorter (about 50%) in a non-final position than in a final one. Notice again that her finding does not just concern Tone 3. After all, the finding seems to point to a natural and common phenomenon in human languages that other things being equal (i.e. factors such as stress and intonation being controled), segments at least tend to be longer in utterance-final position than when preceding another segment of any shape.

Hence I conclude here that if one should regard Tone 3 as significantly longer (i.e. being reflected in phonological terms) in utterance-final position than elsewhere, one should also treat the rest of the Mandarin tones in the same manner. Now, let us come to the second implicit assumption, made in the previous studies mentioned earlier, that Tone 3 is significantly longer than the rest of the phonemic tones when they occur in utterance final position. I will show that findings from certain experimental studies fail to support this claim.

4.2.2.1 Howie's Study

In his quite comprehensive *Acoustic Study of Mandarin Vowels and Tones*, Howie (1976) measures acoustically native speakers' utterances of 136 citation syllables which belong to 34 types of syllables. He measures these 136 syllables in terms of their duration, fundamental frequency, and other acoustic factors. The following are some of the results from his study. As what is of concern here is tonal duration, I have excerpted from these results only the part concerning duration.

(4.10).

Duration of the voiced part of the syllables in milliseconds (Howie 1976, pp. 155-199)

	syllables	Tone 1	Tone 2	Tone 3	Tone 4
1.	yi	182	190	198	230
2.	yu	180	222	276	220
3.	wu	180	242	162	226
4.	ai	270	250	240	266
5.	ya	242	300	324	270
6.	yao	310	272	358	296
7.	you	290	334	390	340
8.	yang	344	328	320	330
9.	yuan	352	430	368	358
10.	ying	310	366	312	300
11.	wen	280	306	290	332
12.	xu	160	140	178	162
13.	hu	166	178	228	230
14.	shi	150	152	212	144
15.	fa	170	238	210	186
16.	ma	280	340	374	330
17.	na	322	332	366	352
18.	la	300	352	340	300
19.	pao	162	196	214	236
20.	pi	104	140	180	144
21.	tu	134	174	186	200
22.	tuo	168	210	250	184
23.	ke	230	210	210	174
24.	bao	254	232	392	244
25.	bi	152	210	200	190
26.	du	192	230	204	216
27.	duo	180	230	242	284
28.	ge	226	210	248	218
29.	qie	134	212	234	220
30.	can	232	232	268	240
31.	chi	210	154	250	150
32.	jie	220	250	232	200
33.	zan	224	350	306	250
34.	zhi	168	210	170	176

A close look at this substantial list of measurements should prompt one to ask: on what basis is the claim made that Tone 3 is significantly longer (by 50%, as Woo claims) than the rest of the Mandarin tones in final position? True, with certain syllable types, Tone 3 does stand out as longer than the rest. An example of such an extreme kind of syllable type is found in the case of No. 24, (and No. 24 only !), *bao*, where Tone 3 is 54% longer than Tone 1, (Tone 1 being the next long-

est after Tone 3). However, such a syllable type is not typically found across the 34 syllable types. Most Tone 3s that are longer than the rest of the tones are much less longer than this percentage (i.e. 54%). There in fact exist 13 syllable types (representing nearly 40% of all the syllable types) that are shorter with Tone 3 than with the rest of the tones. Examples of such a syllable type are found in No. 3, 9, 10, 26, 33, 34, to sample a few in the above list. If one assumed that Tone 3 is 50% (or even 30%) longer than the other tones, one would have to exclude these 13 cases from consideration.

It seems clear that if the choice of the syllable types is non-representative, the result from an experiment based on these non-representative types will surely come out biased. From this, it follows that one cannot make a claim about which of the four tones is significantly longer than the rest on the sole basis of one or even a handful of syllable types. Unfortunately, however, such an approach seems to have been used in phonological studies of Mandarin. An example of the use of such an approach is found in Woo (1972), which will be addressed in the next section (i.e § 4.2.2.2).

Since the claim that Tone 3 is longer than the rest of the tones when they occur in a final position is found to be true only with a subset (about 50% of all the 34 syllable types), it appears rather inappropriate to maintain the claim and to take it as a generalization for all the syllable types involved.

Another interesting observation that can be made, based on Howie's measurements, is that the other tones may also last longer than the rest of the tones on a given syllable. To name a few: Case 19 in the above list shows that Tone 4 is 45% percent longer than Tone 1, and 20% longer than Tone 2. In Case 33, Tone 2 is found to be 56% longer than Tone 1, and 40% longer than Tone 4. Case 1 shows that Tone 4 is 26% longer than Tone 1, 21% longer than Tone 2, and 16% longer than Tone 3. The list could be extended.

The point I wish to make is that if one were to represent Tone 3 with 1/3 more length than the rest of the tones based on cases such as No.24 in the list, one should also take the cases just mentioned into account and represent them accordingly. For instance, the representation should also show that Tone 4 is longer than Tone 1 (based on Case 19), or, Tone 2 is longer than Tone 1 and Tone 4 (based on Case 33), etc. This is already beyond the scope of a phonological description. Thus, even if Tone 3 is found longest among the four tones in some of the cases, the difference may be due to low level phonetic or acoustic details which have no place in a phonological representation.

So far, I have been comparing Howie's syllables more or less on an individual basis. Here, one may argue that some claim can still be made about Tone 3 being longer if it is proven longer collectively on an average. First of all, it should be pointed out that a phonological generalization is not normally made based on a average score. If indeed Tone 3 is longer on the average than the rest in final positions, the most one can claim is that Tone 3 TENDS to be longer than the rest, and the claim is by no means a generalization.

Secondly, even if one did consider the average score in the phonological theory, there would still be no reason to maintain the "superior" status of Tone 3 as previously held. The following table contains the average scores of the duration of each of the four tones and the percentage value by which one tone (in column) is longer than another (in row). The table is a direct calculation from Howie's data as given in (4.10).[5]

(4.11).

	Tone 1	Tone 2	Tone 3	Tone 4
	220	248	263	241
Tone 2	12%			3%
Tone 3	19%	6%		9%
Tone 4	9%			

The percentage scores in the cells indicate how much longer a tone in the left column is than a tone in the top row. What this above table reveals is that the average duration across a pool of 34 syllable types or 136 syllables (i.e. 34 X 4) is 220, 248, 263, and 241 msec for Tones 1, 2, 3 and 4 respectively. On the average, Tone 3 is indeed the longest, with Tone 2 being the next longest, followed by Tone 4. Among the four, Tone 1 turns out to be the shortest.

However, although Tone 3 is the longest, it is only up to 19% (rather than 50% as Woo claims) longer than any of the tones. Furthermore, while Tone 3 is 19% longer than Tone 1, the percentage gets much less when tone 3 is compared with Tone 2 and Tone 4, where the percentages are a mere 6% and 9% respectively. Nowhere in the table is the percentage of 50% or even 30%, as indicated in previous analyses, or one that is close to either seen. These figures considered, it would seem questionable to give the "superior" status to Tone 3 in length as has been done in previous analyses, and the low percentage (i.e. 19% or less) makes one wonder whether the difference in length has any phonological significance or merely low level phonetic implication having no phonological consequences.

Furthermore, if one should believe that the 19% extra length should be formalised in phonological terms, then it should not be unreasonable to use comparable means to encode the extra 12% by which Tone 2 is longer than Tone 1, or the 9% by which Tone 4 is longer than Tone 1, and so forth. However, if one does so, the whole system of tonal representation would be thrown into total chaos; it would be lost in phonetic details so that no generalization could be captured and no insights could be effected.

4.2.2.2 Woo's Study

If Howie's study has not found Tone 3 significantly longer than the rest of the tones in isolation (or in final position), one may want to ask where that figure of 50% comes from, and on what basis it is claimed. I find that the figure is due to Woo (1972), and based on the following measurements made by her (adapted[6] from Woo, Table 2.1, pp. 28-29)

(4.12). Citation Tones of Mandarin Syllables:
Duration Measurements of Components in csec.

	Total Syllable	Initial Cons. (+Glide)	Vowel	Final
I. Tone 1				
ba	36	2	34	
ban	38	2	21	15
bang	38	2	21	15
bei	35	2	33	
II. Tone 2				
ma	42	6	36	
mai	44	8	36	
man	48	6	23	19
mang	44	6	23	15
mao	40	8	32	
mei	38	4	34	
men	42	4	17	21
meng	42	6	19	17
mi	38	4	34	
mian	48	6	20	14
min	36	2	20	14
miao	42	4	32	
ming	42	6	16	20
mou	44	8	36	
III. Tone 3				
da	53	2	51	
dai	52	2	50	
dang	52	2	35	15
dao	59	2	57	
dei	52	2	50	
deng	53	2	37	14
di	57	2	55	
diao	57	2	51	
ding	53	2	36	
du	55	2	53	
duan	61	2	40	15
IV. Tone 4				
zha	31	0	31	
zhai	29	0	29	
zhan	27	0	27 (V+C)	
zhang	23	0	23 (V+C)	
zhao	25	0	25	

The measurements listed in the above table show indeed that Tone 3 is much longer than the other tones, by in fact 50% or more on an average. However, a close look at the table reveals that it is not clear whether the duration difference is truly caused by the difference in tonal categories. For one thing, it may be attributed to the difference in the initial consonants.

Notice that in the above measurements of Woo's, each of the four tones is measured on syllables with a different initial - Tone 1 is measured on syllables with an initial <b> (or /p/ in IPA), Tone 2 is measured on syllables with an initial <m> (/m/), Tone 3 is measured on syllables with an initial <d> (/t/), while Tone 4 on syllables with an initial <zh> (/ž/) - rather than all of the four tones being measured against the same syllable type with the same initial. Given the fact that different consonants may have different effect on the duration of their adjacent segments, the above comparison of tonal length has not been made on an equal footing.

Woo does show her consideration of the difference in the initial consonants by making separate measurements for the duration of the consonant and that of the following vowel. For the syllable dǎ, for example, the initial consonant is measured at 2 csec., while the vowel at 51 csec. This way, she can compare the length of the syllables minus their initials across tonal categories and claim that the difference in the initials is made unimportant. However, this measure still will not make the comparison right. It still will not make the difference in the initials unimportant, since what matters is not simply the absolute duration value of the initial consonant, but the varied influence the consonants may have over the duration of the segments that follow. Compared with Woo, it is clear that Howie's measurements provide a more satisfactory basis for comparison of tonal duration across tonal categories.

Based on the discussions made in the last two sections (i.e. of Howie and Woo), I conclude that Tone 3 has not been found significantly or consistently longer in any context than the other three Mandarin tones. Let me re-capitulate the two arguments I have made in this section (i.e. § 4.2.2) as follows:

(4.13). a. All Mandarin tones are longer, just as Tone 3 has been observed to be, in the utterance-final position than elsewhere.

b. Tone 3 has not been found significantly longer, in the utterance-final position, than the rest of the tones in the same position.

4.2.3 The Tonemic Representation of a Tone

To put the arguments advanced in the previous section into technical terms, I propose the following tonemic representation of the four Mandarin tones:

(4.14).

Tone1	Tone2	Tone3	Tone4	
HHH	MHH	LLM	HML	(in utterance-final position)
HH	MH	LL	HM	
		MH		

This analysis stipulates that a tone contains three tonemes when there is no other tone following, but two when there is. The crucial feature of the analysis is that every Mandarin tone is seen as having an extra toneme when appearing in final position. As I have already shown, this characteristic marks one of the departures of our analysis from traditional ones where Tone 3 alone is thus treated.

It is interesting to note that the above analysis actually explains why it is that only Tone 3 has been thought (or perceived) to be longer in a final position than elsewhere. The reason is rather simple: the extension in Tone 3 is marked by an abrupt change in the direction of the pitch - the third toneme abruptly turns the pitch of its preceding toneme L upwards, making the extension perceptually more observable. On the other hand, in none of the other three tones is the extension likewise marked. In the case of Tones 1 and 2, for example, the pitch level simply

continues from the pitch of the second of the three tonemes. In neither is a change of pitch direction, as seen in the case of Tone 3, detected. In the case of Tone 4, the direction of the pitch is falling and the falling simply continues naturally while there is still temporal space for it. In short, none of the three tones of Tone 1, 2 and 4 changes its fundamental frequency abruptly at the third toneme.

Without an abrupt change in fundamental frequency, the only difference between the longer three-toneme form and its shorter two-toneme alternant becomes, in a sense, a matter of duration. Recall the finding made in the research study by Howie (1976), which was mentioned in § 3.3. He finds that duration does not affect a native Mandarin speaker's judgement of tonal categories. This finding suggests here that a shorter tone HH will be judged to be the same tone as a longer HHH; likewise, a tone of the form HM could be judged to be the same tone as its longer extended form HML.

Notice that the tonemic representations for Tones 1, 2 and 4 have a perfect correspondence with their respective traditional scaler values in terms of the coverage of pitch range. Let me take Tone 4 to illustrate the point: the pitch range covered by the tonemic representation of HML is identical to that covered by the scaler representation of (51). A little explanation, however, is necessary here for the representation of Tone 3 with LLM rather than MLM, though there has been little controversy in the literature with regards to the use of the former. This is because, given the scaler value of (213) for Tone 3, the latter is just as logical a candidate for the representation of Tone 3 as the former.

Hockett (1947) is probably the first person to describe Tone 3 without its rise as a low flat tone. Later, in a Chinese dictionary compiled by Chao and Yang (1947), Tone 3 is described as a low flat segment attached to a slight fall at the beginning, before starting its final rise. According to them, the flat segment lasts

about half of the time afforded a complete Tone 3. They therefore call Tone 3 without the rise "half third". Treating this "half third" as phonologically a flat low tone receives explicit support from Woo, who gives the following explanation for the treatment, based on her spectrographic result:

> "Although the contour [for Tone 3] is traditionally described as `falling-rising', it is not accurate to say that the contour is one which could be graphically represented as \/. The frequency curve does not fall throughout half of the sonorant duration, nor is the fall as sharp as the rise. If one looks at the spectrographic evidence, one will see that the actual contour shows a very slight drop at the beginning of the sonorant articulation, followed by a level low fundamental frequency curve... We do not feel that the initial drop is distinctive; we feel that it is the natural consequence of the articulatory mechanism involved in producing the following low level tone." (p.43)

I think Woo's explanation is essentially correct. From the point of view of articulatory phonetics, it is quite possible that the low tone, being at the lowest pitch of the voice, is hard to reach right away. Thus, even if the speaker's aim (i.e. what his phonological grammar tells him) is to start on a low tone (at the pitch level of 1 on Chao's scale) when articulating Tone 3, he still has to experience a transitional stage of a slight pitch drop at the beginning of the articulation. This slight drop, however, has no phonological significance.

4.2.4 Mandarin Evidence for Treatment of Contours as Units

For the last three sections, I have been defending my analysis proposed in § 4.2 at the tonemic level. In this section, I will move a level up and address the T or tonal level in this analysis. In the literature review given in Chapter 2, I have shown that in two recent studies, Yip (1989a,b) gives evidence from the Wu dialects to demonstrate that Chinese tones, including contour tones, should be treated as units at some level of the tonal representation. As is shown below, evidence for such a treatment has also been found in Mandarin.

While all the major tone sandhi processes in Mandarin are purely phonetic or phonological without any interaction with morphology, there is a special kind of tone sandhi process in this language that affects two very frequently used morphemes in the following manners:[7]

(4.15)

a. yi

in final position:	shí yī	"eleven"
	dì yī	"first"
	yī	"one"
before Tone 1:	yì tiān	"one day"
	yì gēn	"one thread"
before Tone 2:	yì tóu	"one (an animal)"
	yì tiáo	"one (thing & long)unit (of)"
before Tone 3:	yì chǐ	"one Chinese inch"
	yì běn	"one copy (of a book)"
before Tone 4:	yí jiàn	"one piece"
	yí qiè	"everything"

b. bu

before Tone 1:	bù duō	"not much"
	bù gāo	"not high"
before Tone 2:	bù qiáng	"not strong"
	bù xíng	"It won't do."
before Tone 3:	bù hǎo	"not good"
	bù duǎn	"not short"
before Tone 4:	bú duì	"not right"
	bú shì	"be not"

The sandhi processes that are involved in these data can be roughly captured in the following two rules:

(4.16). a. Tone 1 → Tone 4 / ___ {Tone 1, Tone 2, Tone 3}
b. Tone 4 → Tone 2 / ___ {Tone 4}

Or, in terms of values of these tones, it can be stated as follows:

(4.17). a. (55) → (51) / ___ {55, 35, 213}
b. (51) → (35) / ___ {51}

Specifically, the morphemes that are affected by these two rules, that is, the negative morpheme *bū* "not" and the number *yī* "one," are pronounced with the first tone when in isolation or in final position,[8] but alternate between Tone 2 and Tone 4 when appearing before other tones.

The sandhi process that is of interest here is shown by the second of the two rules: the one that changes the rising Tone 2 to the falling Tone 4. The complete change of direction of the tonal contour in the process described by this rule makes it very difficult to account for the process at the tonemic level, in terms of a change of the feature geometry at that level. Suppose one represents the two alternants of the affected tone as LH and HL respectively - to make a simple illustration, it is very difficult to provide any reasonable account to explain why L has become H or H has become L. Even if some bizarre rule could possibly be devised for the change, one would still miss the nature of the change. On the other hand, the nature of this change is obvious if one sees the process at one level above. There, it becomes clear that the process is a dissimilative one which turns a falling contour to the opposite direction when it is before another falling contour. This means that it takes the following tone as a whole unit to trigger the preceding tone to alter as a whole unit.[9]

4.2.5 A Theory of Mandarin Timing Units

An integral part of the tonal representation proposed in this study is the theory of timing units. In this section, I will spell out this theory in a more formal way. Essentially, the theory consists of the following equation:

(4.18). τ(toneme) = 1/3 τ(syllable in citation)
where τ = time

This rule says that a toneme lasts, in a phonological sense, 1/3 of the time of a syllable. Given that a Mandarin tone is co-extensive with the syllable, (4.18) can be extended to the following:

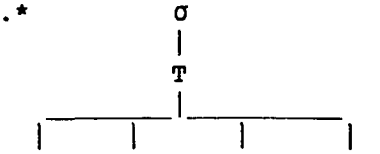
(4.19). A definition of the Mandarin Timing Units

τ(toneme)
= 1/3 τ(syllable in citation)
= 1/3 τ(tone in citation)

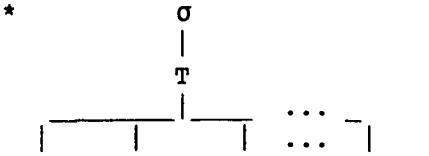
where τ = time

This rule says that a toneme constitutes, in a phonological sense, 1/3 of the time of a tone or a syllable. One of the immediately noticeable merits of this equation is that it can serve as a constraint which effectively excludes an ungrammatical representation of more than three tonemes such as the following:

(4.20). a.*



b.*



where tm = toneme

Both (a) and (b) in the above contain more than three tonemes, and therefore, violate the rule in (4.14). For instance, in (a), a special case of (b), the period of

time represented by the four tonemes exceeds the time allotted to the syllable, in violation of (4.14), and is thus correctly ruled out. Thus, with the timing concept as defined in (4.14), the confinement of the occurrence of tonemes to a number of 3 within any given syllable or tone is achieved.

The reason to limit the number of tonemes to 3 is due to the well-known fact that a tone (not a melody) has not been found to be larger than three tonemes (or contain more than two basic contours of rising and falling) in any language. The most complex tones that have been found are the convex or concave tones which can be sufficiently described in terms of a combination of three (and no more) tonemes such as HLH and HMH. In certain cases, one may occasionally find a tone with four tonemes. These, however, are always the result of a tone-intonation interaction at a low phonetic level, which goes beyond the area of representation at the phonemic-tone level.

However, one of the problems in nearly all of the previous studies of Chinese tones which decompose the tone into its elements is that the occurrence of the number of tonemes is not properly constrained. In these studies, it is theoretically possible for a tone to contain any number of tonemes, and, given the timing function of the tonemes, to go to any length.

For instance, Chao's scale of five theoretically allows a tone to have the shape of, say, 353535, or even longer, though doubtlessly, such a tone or the like never occurs in any human language. As well, nothing in Yip (1980a, 1989a,b) would prevent a tone such as LHLHLH or even longer from happening. These analyses of Chao and Yip clearly suffer from serious over-generalization, as they can potentially generalize numerous tones not substantiated in natural languages.

As already discussed in the literature review chapter, one study that does try to limit the occurrence of tonemes to a certain number (two in this case) for any

given tone is the one by Bao (1990). Pointing out that Yip's (1989b) analysis of Chinese tones theoretically allows more than three tonemes (see § 2.9) to occur in a row, Bao imposes on his analysis a constraint which states that "underlyingly, a contour node may have at most two branches (p. 61)." This confines the number of tonemes for each tone to two.[10] Notice that the use of this constraint is case-specific; no other independent motivation exists behind its stipulation, and it says nothing as to why only two branches are permitted.

Compared with Bao's approach, the rule in (4.14) is more motivated, as it not only highlights the temporal correspondence between the syllable/tone and its tonemes, but also, without further stipulation, predicts that the number of tonemes per syllable/tone can be no more than three for any given Mandarin syllable/tone. In other words, the limitation of the number of tonemes falls right out from the theory of timing relation stipulated by the above equation. Also, with its tonemes empowered by the equation, the present theory is in fact challenging the traditional weight theory. In view of the large scope of this issue, the following chapter is devoted entirely to its discussion.

4.3 Notes to Chapter 4

1 There are in fact several possible ways to represent this tone, depending on the theoretical framework adopted (cf § 2.7, § 2.8, and § 2.9). However, this fact should have no bearing on the point I am making here.

2 In fact, this system of terminology was used earlier in the literature review, though in an implicit and unsystematic way. It is also to be noted that definitions in this system are adopted here as working definitions only, adopted simply to facilitate the development of the present analysis. I do not claim that this is the best system of tonal terminology for the analysis of the Chinese tones.

3 Or, Halle and Stevens' (1971) [+/-stiff] and [+/-slack] will also suffice.

4 See the next chapter for a discussion of "quantity-sensitive" as opposed to "quantity-insensitive" languages.

5 Another experimental study of the Mandarin syllables by Howie (1974) has quite similar results. Though the absolute values of the duration of the syllables are slightly larger than the corresponding ones in his 1976 study, the percentage by which one tone is longer than another remains the same - as shown below.

	Tone 1	Tone 2	Tone 3	Tone 4
	224	253	268	245
Tone 2	12%			3%
Tone 3	19%	6%		9%
Tone 4	9%			

Since the results from the two studies come out almost identical, I will just use Howie's 1976 study for our discussion of the tonal duration.

6 Woo's system of transcriptions of the Mandarin words is replaced with *pinyin* transcriptions.

7 The following data are from the book *Xiàndài Hànyǔ* (Modern Chinese), compiled by four teacher's colleges in China in 1978, the first in the list being *Yāntái Shīzhuān* (Yantai Teacher's College).

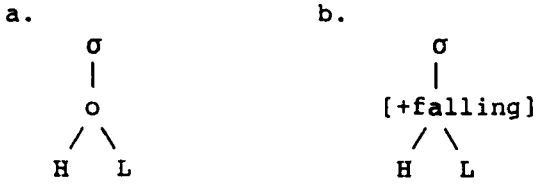
8 The underlying tone for *bū* may be debatable. In dictionaries published in mainland China, it is listed as a fourth-toned morpheme *bù*. This is probably due to the consideration that this morpheme is never in the first tone when appearing before other tones. This fact, however, may in turn be attributed to the fact that this morpheme never appears in final position and almost never appears in isolation. But, when it does appear in isolation, in such situations as seen in the following sentences, it is often heard pronounced with the first tone (irrelevant tonal information is not provided below):

"bū" zhege zi you si hua.
not this word have four stroke
"The word/character "not" contains four strokes."

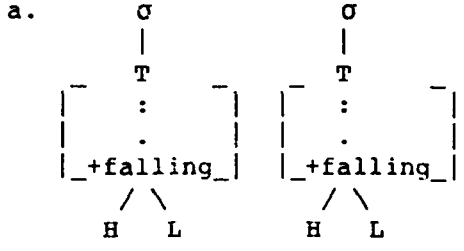
```
Dai Qing you  pian xiaoshuo ming jiao <<Bū>>.
Dai Qing have CLS  novel    name call "not"
"Dai Qing has written a novel named 'Not'."

where CLS = classifier
```

Thus, it does not seem totally unreasonable to assume that it is in Tone 1 in isolation. It should be noted that this controversy is trivial in the present discussion. As will be shown soon, the rule that is of real concern here is (4.17b) and not (4.17a).

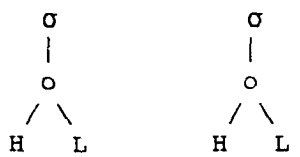
9 The preceding discussion brings up another complicated issue concerning the representation of tones. This is the issue that has been discussed earlier in the literature review: whether to allow contour features, as first developed in Wang (1967), in the system of tonal representation. Yip's (1989a,b) position is that although one should recognize the unitary status of contour tones, yet all one has to do is to recognize a unitary node without the use of contour features. With Tone 4 as an example, the difference between the two approaches can be shown:



According to Yip, (a) in the above is accepted, but (b) is not. However, the Mandarin evidence just discussed seems to suggest the presence of a contour feature such as [+falling] in the two tones under discussion. This is shown in (a) in the following:



b.



Representing the two tones as in (a), one is then able to explain the dissimilation in a rule which may look like the following:

[+falling] → [-falling] / ___ [+falling]
or, [+falling] → [+rising] / ___ [+falling]

Although it is obvious that the adoption of contour features such as these will make the grammar more powerful, it is not clear to me how one can account for this dissimilation process in a more straight forward way. However, I will not pursue this problem any further, as such a pursuit would go beyond the scope of this dissertation.

10 As discussed in the literature review, Bao's clause does not perform what it is meant to perform without further stipulating that the two branches in question can branch no further (see also Note 9 of Chapter 2).

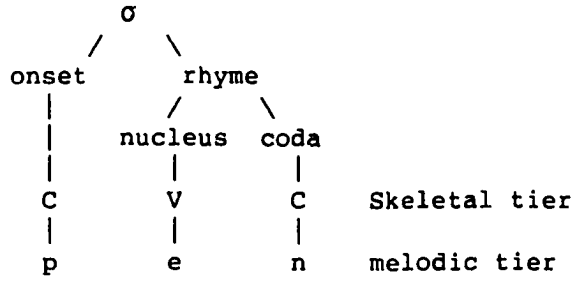
Chapter V

THE MANDARIN SYLLABLE

5.1 *Preliminaries*

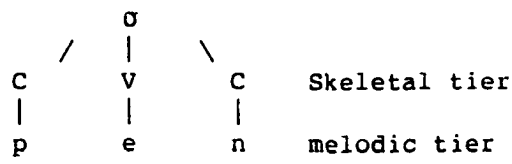
Before I start the discussion of the Mandarin syllable, let us first have a look at the different approaches that have been proposed for syllable analysis. Broadly speaking, there exist two competing analyses: the hierarchical and the flat analyses. The major difference between the two is that the former (the traditional one of the two) allows for an intermediate level or levels between the CV skeletal tier and the syllable tier while the latter does not. For instance, the English syllable *pen* has the following structure in the hierarchical analysis (Kiparsky 1979, McCarthy 1979 and Harris 1983 based on the earlier analysis of the structural school Hockett 1955, Pike 1967).

(5.1).



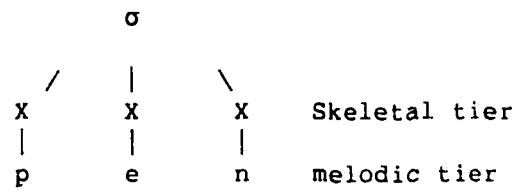
As opposed to the hierarchical structure, a flat structure is proposed by Clements and Keyser (1983) in which the skeletal tier is linked directly to the syllable:

(5.2).

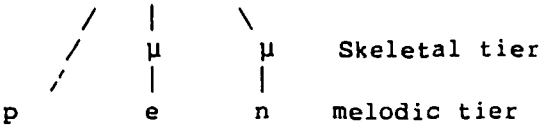


More recently, this flat structure has led to two other alternative flat structures being developed, the X-theory (Levin 1985, Lowenstamm and Kaye 1985) and the Moraic theory (Hyman 1985, McCarthy and Prince 1986, 1990). Both use other constituents in place of Cs and Vs. The X-theory uses unspecified Xs (i.e. variables that may be either Cs or Vs) as the constituents on the skeletal tier, while the Moraic theory adopts the moras. The following illustrate these two approaches:

(5.3). a.



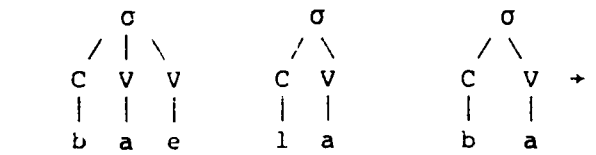
b.



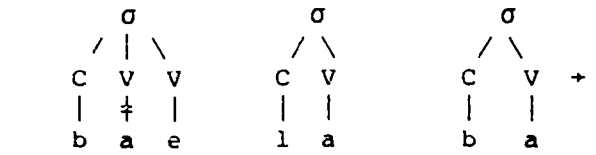
Each of these approaches is supported by a host of arguments based on data from various languages. However, what is of concern here[1] is one major assumption that is shared by all of these analyses. This is the assumption, first proposed by McCarthy (1979) and fully developed by Clements and Keyser (1983), that an abstract skeletal tier of timing units should be adopted.

One of the most important arguments for this abstract tier comes from the phenomenon of compensatory lengthening, observed in many languages. In Luganda (Katamba 1989, pp.171-2), for instance, after morphological concatenation, the sequence /ba-e-lab-a/ becomes [be:laba], with the deletion of the first /a/ and the lengthening of /e/. This case is readily accountable if an abstract skeletal tier is assumed.

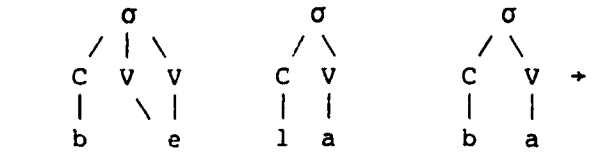
(5.4).



a. deletion/delinking



b. right-to-left spreading



[be:laba]

After the deletion, spreading obligatory occurs to link the only adjacent vowel segment in the syllable to the empty V slot. Given the timing function of the unit of V, the segment /e/ that gets linked to it is lengthened.

The positing of the abstract timing (or weight) tier also allows one to characterize syllables in terms of their binary feature of weight, namely, whether they are *heavy* or *light*. A syllable with more than one V (or μ in the Moraic theory) in the rhyme (i.e. roughly speaking, a syllable minus the initial consonant) is a heavy syllable; otherwise, it is a light syllable.

Phonologists are interested in measuring the weight of the syllable because it has been very frequently found to affect or effect phonological processes. In certain languages, for instance, the weight of the syllable determines stress assignment, and it is normally the heavy syllable that attracts stress. An example of such a language is Alaskan Yup'ik (Krauss 1985), in which all syllables with long vowels or diphthongs are stressed.

A language that has phonological rule or rules sensitive to syllable weight has traditionally been called *a quantity-sensitive* language, while one that does not

have such a rule has been called a *quantity-insensitive* language. Traditionally, Mandarin has been classified as belonging to the latter group along with all other Chinese languages. This classification implies two underlying assumptions. One, Mandarin syllables may vary in weight (or quantity). Two, none of its phonological processes, however, pay attention to its syllable weight.

As I will demonstrate below, these assumptions are erroneous. I will argue instead that all Mandarin syllables have the same weight (or quantity) as far as their internal structures are concerned. Therefore, rather than being considered a quantity-insensitive language, Mandarin should be more appropriately called a *quantity-constant* language. Being quantity-constant is the reason behind the observation that no Mandarin phonological rules have been found making reference to its syllable quantity. Without a change in the quantity, it is only expected that no quantity-sensitive rules are found. In the following section, I will give evidence to show that Mandarin syllables are quantity-constant.

5.2 The Quantity of the Mandarin Syllable

The Mandarin syllable has a canonical structure as in (5.5), which means that it may contain from one (the peak vowel) to four components:[2]

(5.5) (C)(G)V{(G),(C)}

where G= glide (or semi-vowel)

That the Mandarin syllable is quantity-constant can be argued from several perspectives. First, let us see some language-external evidence from errors made by Mandarin English-as-a-second-language (ESL) speakers.

One of the most important characteristics of a quantity-sensitive language is that it makes a contrast between long and short vowels with or without a contrast between long and short consonants. Katamba (1989) shows, for instance, that lan-

guages in which a distinction is made between light and heavy syllables fall into two groups. One has light syllables such as /a/ and /pa/ and heavy syllables such as /paa/ and /pat/. The other contains light syllables of the type /a/ and /pat/ with heavy syllables of the type /paa/ and /paat/. In both groups, a short vowel is contrasted with a long vowel. Among languages of the first group, it is vowels in /pa/ and /paa/ that are contrasted, whereas those of the second type, it is vowels in the pairs of /pa/-/paa/ and /pat/-/paat/ that are contrasted. In contrast to these languages, Mandarin lacks totally such a distinction between vowels or consonants. This is often shown in the errors made by Mandarin speakers in their pronunciation of English words. Very frequently, for instance, many of these speakers substitute a long vowel /u/ as in "food" for a short vowel /U/ as in "book", a short vowel /I/ in "bit" for a longer vowel /i/ as in "beat," and a short vowel /ɔ/ for a long vowel /ɔ:/ (British English) in "court".

Not only does Mandarin lack a quantity contrast among vowels, but it also lacks such a contrast between single vowels and diphthongs. In normal speech in Mandarin, diphthongs are not necessarily longer than monophthongs, as they are in a quantity-sensitive language. To a native English speaker, the Mandarin diphthongs in normal speech may sound like monophthongs. As Kratochvil (1968) observes,

> "[Mandarin] diphthongs and triphthongs are complex vowels, not combinations of individual simple vowels...Generally speaking, [they] have slightly greater duration than simple vowels, but much smaller than the combination of the duration of their components occurring as simple vowels." (pp. 30-31)

He gives the following examples to illustrate the difference:

> "The duration of vowels in the syllables [pà] bà, [pò] bò, and [pào] bào recorded by the same [Mandarin] speaker in similar conditions was measured with these results: [-à] in [pà] 0.23 sec, [-ò] in [pò] o.22 sec, and [-ào] in [pào] 0.28 sec. [-ào] lasted longer than either [-à] or [-ò] (0.28:0.23, 0.22 sec), but its duration was much smaller than the total of durations of [-à] and [-ò] (0.28 sec: * 0.45sec)." (pp. 31)

It seems clear that the slightly longer duration (by a mere 0.04 or 0.05 sec) of the diphthong [-ào] (0.28 sec.) as opposed to the duration of the single vowel [-à] (0.23 sec.) or [-ò] (o.22 sec.) is not phonological.

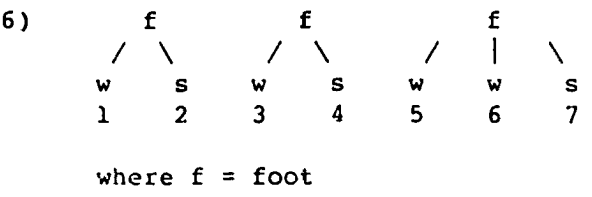
The lack of quantity contrast between Mandarin diphthongs and single vowels provides good explanation for why native Mandarin speakers are often heard replacing short /I/ with the diphthong /ei/ when they pronouncing English words such as "s/I/t down" and "c/I/vil war." These become "s/ei/t down" and "c/ei/vil war" respectively.

Quantity contrast is also found lacking between the native Mandarin speakers' pronunciation of the English syllables of the shape CV and those of the shape CVR (R being liquids /r/ and /l/, or [-anterior] nasals /n/ and /ŋ/). For instance, native Mandarin speakers tend to say /ka/ for /kar/ "car", and /sĩ/ for /sin/ "sin". Native Mandarin speakers are also found to shorten a longer syliable in other ways. Of the dozens of Mandarin speaking students I have checked with (all are from mainland China and studying at the University of Victoria), none is able to pronounce the name Brian correctly. Instead of /'brayən/, they usually say simply /bran/.

The above evidence clearly indicates that native Mandarin speakers do not judge the length of a syllable by the number of consonants or vowels in it. It seems, rather, that they have in their grammar a quantity template at the syllable level, and no matter how many components there are in a syllable (whether it is a syllable in their own mother tongue or one of a second language), they will tend to "shrink" (or "stretch") them to fit into that template.

That the Mandarin syllable is quantity-constant is also supported by evidence from Mandarin (or Chinese in general) versification and the related phenomenon of stress assignment. In the versification of a quantity-sensitive language, it is normally the case that the formation of, say, an iambic foot requires that the first syllable be light while the second heavy. The reverse is true for a trochaic foot. In Latin, for example, a word such as *carō*[3] "flesh" with a light syllable CV and a heavy syllable CVV is an iambic foot, while a word such as *mēnsa* "table" with a sequence of a heavy and a light syllables can potentially be a trochee. On the other hand, the same word *carō* cannot serve as a trochaic foot neither is *mē* iambic. In contrast, all Mandarin syllables, no matter how many components they contain, can potentially be in any of the syllable positions in an iambus or a trochee, or any type of foot at all. For instance, in a Chinese regulated verse (or *lüshi)* with seven syllables in each line,[4] the metric structure is as follows (cf. Chen 1979, Yip 1980b):

(5.6)



where f = foot

This metrical structure contains two iambic feet followed by an anapaest[5]. The point of concern here is that unlike the syllables in Latin metrics, any Mandarin syllable, no matter what its internal structure, may potentially be in either the weak or the strong positions of these feet.

Another clear example of such a free distribution of Mandarin syllables involves a dissyllabic word whose second syllable is in the stressless neutral tone (it is, therefore, a trochee). In such a word, the second syllable is unstressed and may last 50% less in duration than the first syllable. Yet, there is no restriction whatsoever as to which Mandarin syllable may fall into that unstressed position.

Free syllable distribution can also be observed in stress assignment in Mandarin. As mentioned before, in quantity-sensitive languages, stress is often found attracted to certain types of syllabic structure, normally the heaviest one. Mandarin stress, however, is not assigned based on the numbers of Cs and Vs in the syllable, but on other conditions, one of which is the position of the syllable in a morphological or syntactic string. In short, any Mandarin syllable may potentially receive primary stress irrespective of its internal structure.

How then can one interpret such a free distribution of Mandarin syllables in stressed as well as unstressed positions? One plausible explanation is that all Mandarin syllable weigh the same, so that even if there are quantity-sensitive rules (such as rules of the regulated verse and stress-assignment), this sensitivity will not be shown by them.

5.3 The Syllabification of the Mandarin Syllable

The quantity-constant characteristic of the Mandarin syllable can also be shown, in a somewhat indirect way, from facts concerning Mandarin syllabification. Traditionally, syllabification in languages has been observed to occur only at or above the morphological level (i.e. after morphological concatenation), and is characterizable or predictable to a greater or lesser degree by rules or principles. However, Mandarin syllable structure is predetermined: except for a few trivial cases of low level phonetic ambi-syllabification (cf. Kahn 1976),[6] there is no sylla-

bification process at or above the morphological level. Arguments for this position are provided below.

According to Goldsmith (1990), there are three competing approaches to syllabification: the "all-nuclei-first" approach, the "linear-scanning" approach, and the "total-syllabification" approach. The differences among these approaches are not crucial here. What is important is that all of these approaches adopt a general principle: the Onset First Principle (Kahn 1976, Clements and Keyser 1983). This principle says:

(5.7). Onset First Principle

> "a. Syllable initial consonants are maximized to the extent consistent with the syllable structure conditions of the language in question.
>
> b. Subsequently, syllable-final consonants are maximized to the extent consistent with the syllable structure conditions of the language in question." (Clements and Keyser 1983, p.37)

However, this principle, absolutely necessary in all three of the syllabification approaches, fails to work for the Mandarin case. Since all three approaches employ the Onset First Principle, one can demonstrate the failure of these approaches in handling Mandarin by proving that the Onset First Principle cannot account for the data. The following demonstrates this inability of the Onset First Principle with the all-nuclei-first approach.

The all-nuclei-first (henceforth ANF) approach is first developed in Kahn (1976). According to him, syllables are built up from the syllabic element first and then adjoining consonants to these syllables following the Onset First Principle and other relevant rules and principles. An example given by Kahn is the syllabifi-

cation of the English word "Boston" which is done in three steps as shown below (ambi-syllabification details occurring in fast speech are omitted here):

(5.8). a. All nuclei first

```
b ɔ s t ə n
|     |
σ     σ
```

b. Onset first

```
b ɔ s t ə n
\|  \ \|
 σ    σ
```

c. Coda Association

```
b ɔ s t ə n
\|  \ \|/
 σ    σ
```

The ANF approach, effective and adequate as it is with the parsing of an English utterance into permissible syllables, cannot account for the creation of the structure of a Mandarin syllable. Although many of the Mandarin words or expressions such as *shāfā* "sofa, couch" and *shūcài* "vegetables" do follow the Onset First Principle, many others do not. For instance, Mandarin has words such as *pi'ao* "leather coat" with two syllables, representable segmentally in terms of /piau/ (or /piaw/ -- the difference is however irrelevant here), as shown below:

(5.9).

```
/ p  i a  u /
   \/   \/
   σ    σ
```

This syllabification result, however, cannot be correctly attained by the ANF approach with its Onset First Principle.

(5.10). a. nuclei first

```
/p i a u/
   \|/
    σ
```

b. onset first

```
/p i a u/
   \\|/
     σ
```

With this approach, one and only one syllable can be derived for this sequence, as clearly shown above. The interesting thing about the result in the above is that it IS a permissible syllable in Mandarin, except that with such a syllable structure, the sequence no longer means "leather coat" but (with Tone 1) "drift". The /piau/ example gives a good indication that a systematic and generalized syllabification

framework such as the ANF analysis cannot possibly account for the two syllables as shown in (5.9); it cannot make a necessary distinction between a dissyllabic word *pi'ao* /piau/ "leather coat" from a single-syllable word *piao* /piau/ "drift".

The word *pi'ao* "leather coat" is, however, not an isolated case. There are in fact many words in Mandarin that are like it, pronounced as two syllables. The following are just a few more examples:

(5.11). pinyin gloss

li'ang	"Lyon (city in France)
bao'an	"report the case to the security"
yindi'an	"American Indian"
nigu an	"Budhist nunnery"
chu er fan er	"go back on one's words (a four-syllable idiom)"
wei'ai	"cancer of the stomach"
bi'ai	"cancer of the nose"
bei'ai	"sorrow"
bo'ai	"universal fraternity"
jia'ao	"a thinly-padded coat"
jiao'ao	"conceited, proud"
ji'e	"hunger"
yu er	"fishing bait"
bao'en	"pay a debt of gratitude"
bo'en	"Bonn (city in Germany)
hai'ou	"seagull"

Another kind of sequence that defies after-concatenation syllabification, such as performed by the ANF approach, includes words or expressions in the following shape:

(5.12).

pinyin	gloss
wan'an	"good night"
mian'ao	"cotton-padded coat"
jian'ao	"suffer"
en'ai	"conjugal love"
lian'ai	"in love"
chen'ai	"dust"
qin ai de	"dear (someone)"
tian'e	"swan"
ran'er	"however"
lian ou	"lily root"
chang'an	"(name of city in China)"
qing'an	"wish someone good health"
dong ou	"short for E. Europe"
xiong'e	"fierce"
chong'er	"pet, favorite"

In these above cases, the first syllable ends in a nasal,[7] while the second begins with a vowel. Indeed, incorrect syllabification would result if the above were to be syllabified in accordance with the Onset First Principle. Instead of *wan'an* "good night", for example, one would obtain *wa nan*, a sequence without obvious meaning in Mandarin.[8] The generalization that emerges from this investigation is quite striking: no matter what gets linked first, nuclei, onset, or coda, and no matter what the combination of the linking processes, the system will not work for a sizable number of Mandarin syllables.

Further evidence to show that Mandarin syllable structure is predetermined comes from the observation that Mandarin syllables form a closed set. This means that there is a relatively small, fixed number of Mandarin syllables (405 in all if tonal differences are ignored or around 1200 if tonal differences are considered). This is clearly shown from the fact that "even when new terms are borrowed from foreign languages, they are always interpreted in terms of the existing set of syllables (Norman 1988, p. 138)." In fact, new syllables are never added to this closed set. Unlike languages with a generative syllabification process, whereby a syllable can be generated and accepted as a possible though non-occurring syllable

(e.g. "blick" in English), there are no such "accidental gaps" among the set of the possible syllables in Mandarin. If a syllable is not a member of the set, it is then by no means a possible syllable in Mandarin. For instance, the syllable *bou* /bou/ is not occurring in Mandarin. It may be considered an accidental gap, since /-ou/ can be combined with any other Mandarin consonant,[9] and /b-/ with most other Mandarin rhymes. However, this *bou* will not at all be accepted as a Mandarin syllable in any circumstances - not when new words are coined, nor when foreign words are translated.

The conclusion to be drawn from the above discussion is that in Mandarin, syllable structure is determined before morphological concatenation, after which there is no more syllabification process. This is one fundamental nature of the Mandarin syllable. Here, one may perhaps ask: why does syllabification not occur after morphological concatenation in Mandarin, as it does in many other languages of the world? Or, rather, why does syllabification occur at all after concatenation? I believe that syllabification after concatenation is intimately related to quantity. In particular, it occurs to ensure that all the syllables conform to the quantity requirements of the language.[10] Mandarin syllables, being prespecified and standardized in terms of their quantity before concatenation, require no further adjustment when they are strung together to form a larger domain, and as a result, no further syllabification rules are required.

So far, I have tried to argue that all Mandarin syllables weigh the same and are quantity-constant underlyingly. I have cited evidence from errors made by Mandarin ESL speakers, Chinese versification, translation of foreign words into Mandarin and Mandarin syllabification, which are all found to lend support to this argument.

5.4 *The Timing Function of the Mandarin Syllable*

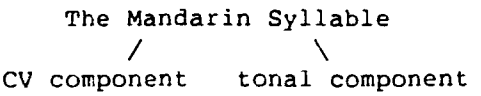
In the traditional weight theory, no matter how syllable weight is represented - by a series of Cs and Vs, as Clements and Keyser (1983) stipulates, or by a tier of moras, as is the practice in Hyman (1985) and McCarthy and Prince (1986), one function that the weight measurement serves is without question to indicate timing. Thus, in a relative phonological sense (as against an absolute phonetic or acoustic sense), syllables that weigh the same are theoretically pronounced with the same amount of time. I have argued in the previous section that all Mandarin syllables have identical weight. Now, a question arises: does this identity mean that Mandarin syllables are always pronounced with the same duration?

I have shown in Chapter 4, Mandarin syllables may very in length depending on their morphological or syntactic positions. In a two-syllable word, for instance, the second syllable is always longer than the first if the second syllable is not in the neutral tone. A similar situation is actually found in a three-syllable word. In such a word, the pattern is usually that the last is the longest followed by the first, and the middle one is usually the shortest (cf. Chapter 6). If, as I have just shown, the quantity of the Mandarin syllable has nothing to do with the actual length of the syllable, what then does? My answer is that the timing function is carried by the tonemes rather than the Cs and Vs of the Mandarin syllable. I shall elaborate further on this point in the following section.

5.5 *The Double Facets of the Mandarin Syllable*

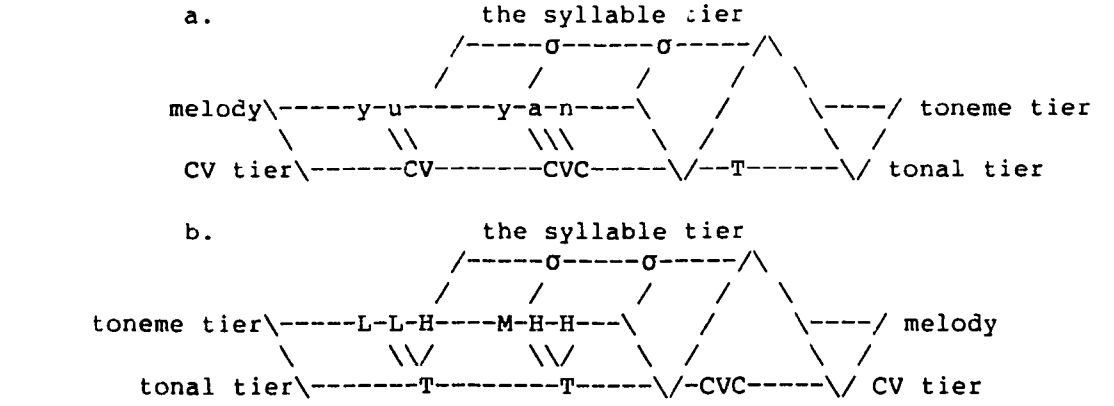
One unique feature of the Mandarin syllable which sets it apart from syllables in many other languages of the world is that it has a tonal tier attached to it in the underlying representation. Thus, besides a CV component, which is universally found in syllables of all languages, the Mandarin syllable has an extra tonal component as shown below:

(5.13).



A three dimensional representation of the same structure is given below with the Mandarin word *yǔyán* "language" as an example:

(5.14).



The two figures above display two viewpoints of the same structure. What is shown by them is that on the one hand (cf. a), the syllable tier is linked to the CV tier which in turn is linked to the melody tier, and on the other hand (cf. b), the syllable tier is also linked to the tonal tier which in turn is linked to the toneme tier (shown in detail by (b)). As mentioned before, many of the languages of the world have only half of the representation, without the tonal part. In these languages, it seems only natural for the CV tier to perform the major (if not all of) the functions (including the timing function) necessary for the syllable to perform. On the other hand, it seems equally natural and logical to see the other side of the syllable, if there is one, "share" its responsibilities. Thus, it is natural to see that in the Mandarin case, the other side, the tonal tier and its tonemic structure, serves the timing function of the syllable. Such a "sharing" may in fact account for the inactivity of the Mandarin quantity-constant CV tier (shown in detail by (a)) in phonological processes.

Here I would like to make a distinction between the two terms of *quantity* and *weight*. If "quantity" is used to refer to the CV (or X in X-theory, or μ in a Moraic theory) make-up of a syllable on the skeletal tier, "weight" can then be used to refer to the timing (or the length) of the syllable. Thus, in a quantity-sensitive language in which there is no autonomous tonal tier, quantity implies weight, and the CV skeletal tier performs the timing function. However, in a quantity-constant language such as Mandarin in which there also exists an autonomous tonal tier, quantity may have nothing to do with weight. Rather, it may be the other way round: weight serves as a template for the quantity as a whole to fit into (cf. § 5.2). The weight itself is determined by the tonal tier, or rather, the tonemic make-up of the tonal tier. Notice that the separate behaviors of quantity and weight may depend on the tonal tier's being autosegmental. If so, there may be yet a third scenario: there may be tonal languages in which tones are properties of the syllabic segments (i.e. they do not form an autonomous tier), and their timing function may consequently still be served by the skeletal tier.

Support for treating Mandarin tonal tier as the timing tier can be drawn further from facts about Mandarin stress assignment and tone sandhi. As I will show in the next Chapter, there is a direct link between stress assignment and the number of tonemes in a syllable, just as stress assignment is related to CV structures in a quantity-sensitive language.

5.6 *Notes to Chapter 5*

1 Unfortunately, it is not possible for me to provide a review of these arguments here, as such a review goes beyond the scope of the present study.

2 For examples of Mandarin syllables, please refer to § 4.2.2.1, § 4.2.2.2, and (5.11) and (5.12) of this chapter.

3 The bar on the vowel signifies a long vowel.

4 Chinese regulated verse is a highly regularized form of versification. It must observe a number of strict rules. These include a fixed number (either seven or five) of syllables in each of the lines, and a fixed number of lines (normally four or eight) within one poem, a designated form of metrics in the lines, a designated form of tone distribution, and syntactic and semantic matching between elements of two lines forming a couplet. (cf. Jiang and He 1987). A very famous couplet from such a poem is given below:

```
chūn cán   dào  sǐ  sī fāng jìn,
là   jù    chén huī lèi shǐ  gān.
1    2     3    4   5   6    7
```

Word for word glossary:
spring silk-worm until die silk then finish
wax fire become ash tear began dry

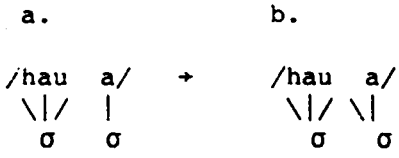
Literary translation:
"A spring silk-worm stops producing silk only when it dies.
A candle-fire extinguishes only when it becomes ashes."

5 The exact shape of the metric structure may be controversial. However, this controversy is of no concern here.

6 For example, it has been observed that the exclamatory particle *a* in Mandarin may optionally receive an onset through assimilation from its preceding segment.

```
hao a  /hau a/  →  /hau wa/
nan a  /nan a/  →  /nan na/
xing a /xiŋ a/  →  /xiŋ ŋa/
hui a  /hui a/  →  /hui ya/
```

The process involved is characterizable in terms of what Kahn calls ambi-syllabification as shown below in (b):



However, such an ambi-syllabification process after morphological concatenation is very rare; it is only used by some native Mandarin speakers; and its

occurrence can be quite inconsistent among these speakers. For instance, it is perfectly acceptable to say

/hau ya/ rather than /hau wa/

7 There are only three consonants that appear syllable finally in Mandarin. The nasals /n/ and /ŋ/ and the retroflex /r/.

8 The existence of words like those in (5.11) and (5.12) has led to the adoption of the convention of *pinyin* writing that an apostrophe must be placed between the two syllables in question.

9 The following are these combinations: *pou*, *mou*, *fou*, *dou*, *tou*, *mou*, *lou*, *zou*, *cou*, *sou*, *zhou*, *chou*, *shou*, *rou*, *gou*, *kou*, or *hou*. These exclude the set of palatals normally listed in the set of Mandarin consonants in the *pinyin* system. However, these palatals *j*, *q*, *x*, or /tç, tçh, ç/ occur only before [+high, -back] vowels (i.e. *i* and *ü*). Historically, they are derived from *g*, *k*, *h*, and are still in complementary distribution with them.

10 Some of these requirements may of course be of a language-universal nature.

Chapter VI

MANDARIN TONE SANDHI

In this chapter, I will provide a unified analysis for Mandarin tone sandhi. I will show that this unified analysis makes it possible to draw insights into the nature of Mandarin tone sandhi. I will also show that this unified analysis will only be possible if the theory of Mandarin tonal representation developed in the last two chapters is adopted. First, let me briefly introduce the Mandarin tone sandhi data. Following traditional practice, these data are provided below in Chao's (1968) theory and notation (which are shown in (i)s in the following).[1]

(6.1). a. Mandarin Fourth Tone Sandhi
(i). 51 → 53 / ___ 51
(ii). Tone 4 (51) → Tone 4 (53) / ___ Tone 4 (51)

b. Mandarin Third Tone Sandhi (A)
(i). 213 → 21 / ___ {55, 35, 51}
(ii). Tone 3 (213) → Tone 3 (21) / ___ {Tone 1, Tone 2, Tone 4}

c. Mandarin Third Tone Sandhi (B)
(i). 213 → 35 / ___ 213
(ii). Tone 3 (213) → Tone 3 (35) / ___ Tone 3 (213)

d. Mandarin Second Tone Sandhi
(i). 35 → 55 / {35, 55} ___ {55, 35, 213, 51}
(ii). Tone 2 → Tone 1 / {Tone 1, Tone 2} ___ {Tone 1, Tone 2, Tone 3, Tone 4}
Note: This last rule applies in fast speech only.

In all, four tone sandhi processes have been observed in Mandarin, and all Mandarin tones except for Tone 1, the only level tone, are found affected by one or the other of these rules. In particular, Tone 3 has three surface values alternat-

ing with one another. They are (213), occurring in utterance final position[2], (35) when Tone 3 appears before another Tone 3, and (21) when it occurs elsewhere. Tone 4 has two manifestations, alternating between the tonal values of (51) in utterance final position and (53) elsewhere. Tone 2 also has two surface forms, (35) in final position, and another form considered identical to that of Tone 1. The latter value of Tone 2, however, occurs only in fast speech. In the following sections, I will discuss and formalize these four tone sandhi processes in detail. I will begin by addressing a very important question involved, as given below.

6.1 Which is Underlying?

In Chapter 4, it is proposed that a Mandarin tone has three tonemes when it appears in an utterance-final position, but two when in a non-final position. It is also implied there that the utterance-final form is the underlying one. In the past, there has been much controversy surrounding the question of which of the two forms should be taken as the underlying form. The traditional position is that the utterance-final form is the underlying one, while the sandhi form is but derived through a sandhi rule. This is the position first proposed by Chao (1930), and later followed by most linguists of Chinese. This is the position obviously shown by (6.1c-i), which is repeated below (a) together with another version of the same rule (b).[3]

(6.2). a. 213 → 21 / ___ {55, 35, 51}
or, b. 3 → ∅ / 21 ___ {55, 35, 51}

The first person to hold an alternative position is Woo (1972). Specifically, Woo reverses the two forms of Tone 3 by putting the form (21) (or LL in her analysis) in the structural description position and considering the form (213) (or LLM[4] in her analysis) as being derived:

(6.3). a. LL (21) → LLM (213) / ___ (pause)
or, b. ∅ → M / LL ___ (pause)

This position is later taken up by Yip (1980a).

The difference between the two approaches can be seen as one between treating the sandhi process as toneme deletion and treating it as toneme addition. I believe that the debate between the two opposite treatments of the sandhi processes is apparently due to the elusive nature of these previous analyses: because of the manner in which the tones are configured in them, the nature of the difference between the two tonal forms is concealed (to be discussed shortly). With the representational scheme I proposed in Chapter 4, the issue becomes trivial. It becomes obvious that the utterance-final tones, or in other words, the three-toneme tones, are underlying. Namely, the traditional approach is considered correct. The reason for this conclusion is simple: if the two-toneme tones were treated as underlying, one would have much difficulty explaining why it is that each of the four tones acquires a different toneme in the utterance-final position. That is to say, why is it that in final position, Tone 3 acquires a M toneme, Tones 1 and 2 a H toneme each, and Tone 4 a L toneme? On the other hand, if three-toneme tones are treated as underlying, the two-toneme tones (i.e. the sandhi forms) are easily explained as the underlying tones losing their final tonemes when they appear before another tone. In short, all cases become explicable by a natural phonetic motivation. This treatment is further defended and defined in the next section.

6.2 *The Nature of Mandarin Tone Sandhi*

Now that it has been determined that the utterance-final forms are underlying, I can proceed to answer the fundamental question of why these sandhi processes occur. My answer consists of the following Tone Reduction Principle[5]:

(6.4). Tone Reduction Principle

In normal speech, reduce a tone by one toneme iff it is immediately followed by another tone within the same prosodic foot.

This principle implies a two-toneme template that operates on any tone that occurs before another tone in normal speech; this principle functions to "tailor" any underlying tone to two tonemes if followed. This principle also implies the assumption that it is the presence of an immediately following tone that is responsible for the occurrences of all these Mandarin tone sandhi processes.[6] As I will demonstrate in the following sections, the adoption of this principle provides a unified account for all the four Mandarin tone-sandhi processes that have been discussed in the literature. I will start with a discussion of the fourth-tone sandhi process.

6.2.1 The Fourth-Tone Sandhi (4TS)

As described at the beginning of this chapter, Mandarin 4TS involves a process in which Tone 4 alternates between the values of (51) and (53) in the following fashion (also see 6.1a).

(6.5). The Fourth-Tone Sandhi (4TS)

51 → 53 / ___ 51

This sandhi process is illustrated below with a few Mandarin examples:

(6.6). a. fàng(51) jià(51) →
fàng(53) jià(51) "have holidays"

b. zuò(51) yè(51) →
zuò(53) yè(51) "homework"

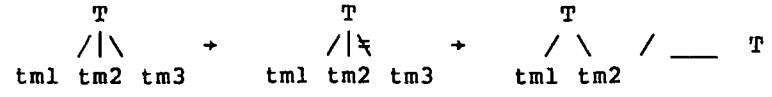
c. zhò(51)ng dì(51) →
zhò(53)ng dì(51) "farming"

d. dì(51) zhè(51)n →
dì(53) zhè(51)n "earthquake"

I argued in Chapter 4 that underlyingly, Tone 4 should, at the tonemic level, be represented as HML, a representation that accommodates both the traditionally-observed pitch coverage (i.e. from 5 to 1) as well as the duration (i.e. three tonemes) of Tone 4. With such a representation, the first of the two juxtaposed Tone 4s must be reduced to two tonemes, in view of the Tone Reduction Principle (henceforth, the Principle).

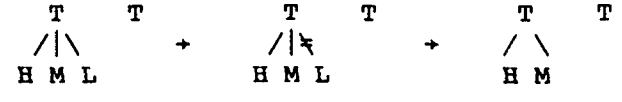
A question hence arises here: how exactly is Tone 4 reduced to HM (rather than, say, ML) by the Principle? It seems that there is an operational rule which serves to implement the Principle, and this rule may be formalized as follows:

(6.7). Tone Truncation Rule (TTR)



What this rule does is, roughly put, to delete the last toneme of a tone when the tone is not in an utterance-final position. It is thus a toneme apocope rule. With this rule, the process described as 4TS falls out. Implementing the Principle, this rule eliminates the L tone of Tone 4, and consequently the right surface result of HM is derived:

(6.8).



In fact, assuming the Principle and the representation of Tone 4 as HML, the analysis just proposed has not only yielded the correct output, but has also brought out insights into the nature of the process. 4TS, under this analysis, is basically a truncation process rather than a tone-replacement process, as it has been configured in previous analyses (see 6.1c). The cause of the process is hence revealed for the first time: the presence of a following tone prevents the affected tone from being realized in full. Thus, the analysis just proposed reveals not only WHAT happens when two Tone 4s are juxtaposed, but also WHY it happens.

Naturalness is yet another unique characteristic of the present analysis, besides its being able to explain rather than simply describe. Segment truncation (or apocope) is a natural linguistic process observed in many languages, so is the motivation for the process to occur (i.e. to satisfy a designated timing requirement). Many languages have compensatory shortening (or lengthening) processes in which a segment, just as Tone 4 in 4TS, is shortened (or lengthened) to fit into a designated timing template.

Lastly, the present analysis actually predicts that Tone 4 has a shorter form not only before another Tone 4, but also before the rest of the tones. This is exactly what has been found in several empirical studies. Shen (1990b), for instance, observes in her acoustic study of Mandarin tones that the pitch of Tone 4 becomes (53) not only before another Tone 4, but also before Tone 1. Dreher and Lee (1966) found, in their acoustical study more than two decades ago, that Mandarin tones including Tone 4 are all 20% shorter before Tone 0. Shih (1987), in addition, has found that a Mandarin vowel, that lasts 300 ms in a final position, lasts only 150 ms when followed, suggesting that a Mandarin tone is considerably shortened when preceding another tone. Shih's study actually expresses explicitly the shortened duration of Tone 4 by saying that, when followed, "a Tone 4 ... never reaches the L target... Tone 4 in the non-sentence final position is actually HM (p.10)."[7] All these findings strongly corroborate the present treatment of 4TS, confirming that Tone 4 is not only shorter before another Tone 4, but also before the rest of the tones.

Compared with the present analysis[8], previous analyses have not achieved these above-mentioned advances. The major problem with them lies in their representational framework for the tones. For example, in Chao's analysis for 4TS in which the two alternants of Tone 4 are represented as (51) and (53) respectively, it

cannot possibly be revealed that the durational discrepancy between the two, shown clearly by our HML and HM, is ultimately the reason behind the alternation. Also, Chao's essentially tone-replacement analysis gives the wrong indications that the derived Tone 4 being (53) in value is purely accidental, and that the result could have been a tone of any other shape. This is obviously a gross over-generalization. The present analysis, by comparison, is much more constrained in this respect, for it only generates HM (53) and no others. Compared with a tone-replacement analysis, the present analysis has another advantage (related to tne one just mentioned) that it can reveal the relationship between the underlying and the derived: the former is but part of the latter.

Now, let me compare the present analysis with Yip's (1980a).[9] This comparison is simple, as Yip has no treatment for 4TS. In particular, she has no representation for the value of (51) (see also Chapters 3 and 4), which means she does not recognize the existence of that value. The question is: is her treatment justified? I believe the answer is no. Recall that Yip (cf. § 2.7) proposes a constraint to curb the generative power of her analysis, and the constraint stipulates that the value of her Register Feature [upper] remains constant within a morpheme (see 3.31). The effect of this constraint is the exclusion from the grammar of any tone that goes across the middle point of the pitch range. Thus, when it comes to the representation for Tone 4 (51), Yip's analysis is in difficulty. This is in fact revealed in her explanation for her representation of Tone 4:

"The choice of [+upper] for the falling tone needs some explanation. Phonetically this tone starts very high (actually above the high level tone) and falls very low: in other words, it covers the entire range of the voice. In the register theory developed here it is necessary to assume that it is phonologically either Upper or Lower Register, and that the extremes of the tone are irrelevant. The reason for assuming that it is Upper Register, and therefore basically /53/ (so that the further fall to [51] is a phonetic detail possible in a language that has only one falling tone) is as follows: in a sequence of two fourth tones the first is realized exactly that: [53] and only the second falls right down to [51]. There is no equivalent phenomenon for any of the other tones, and it is most easily explained by assuming that this is because what is important about this tone is that it is high and falling, rather than how far it falls. Notice that these facts would be quite inexplicable if we took the tone to be [-upper] register, and a special sandhi rule would be required." (p. 183).

The above explanation given by Yip for treating Tone 4 as [+upper] shows her difficulty: she has to be confined to the only two choices that are permitted by her theory (cf. Chapter 3) for the representation of Tone 4: either [+upper] (i.e. (53)) or [-upper] (i.e. (31)). As (31) is quite obviously unacceptable, she virtually has only one choice: (53). Indeed, it is due to the limit imposed on tones by her theory, rather than any other reason, that she has to exclude (51) from the grammar (also see § 3.6).

In fact, Yip's explanation reveals that her position involving the representation of Tone 4 is contradictory. On the one hand, she recognizes the existence of the value of (51) by her observation that "phonetically this tone starts very high (actually above the high level tone) and falls very low: in other words, it covers the entire range of the voice". On the other, she has no phonetic representation for that value. She cites the process of 4TS ("in a sequence of two fourth tones the first is realized exactly that: [53] and only the second falls right down to [51]") to motivate her choice of (53) as the phonological shape of Tone 4, yet does not recognize the existence of this very sandhi process by not having a representation for one of its two variants (i.e. (51)). Clearly, Yip's analysis has adequately handled neither the value of (51) nor the process of the fourth tone sandhi.

6.2.2 The Third-Tone Sandhi: A (3TS(A))

Let us first have another look at the 3TS(A) process configured by Chao (also see 6.1b):

(6.9). Mandarin Third-Tone Sandhi: A (3TS(A))

213 → 21 / ___ {55, 35, 51}

Examples to show this sandhi process are given below:

(6.10). a. lǎo(213) shī(55) →
lǎo(21) shī(55) "teacher"

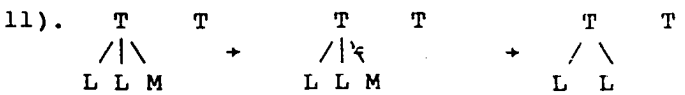
b. lǎo(213) Lín(35) →
lǎo(21) Lín(35) "Old Lin(a surname)"

c. lǎo(213) dà(51) →
lǎo(21) dà(51) "the eldest"

The cause for the 3TS(A) process also originates from the existence of the Principle in (6.4). With this Principle, and with an underlying representation of LLM, Tone 3 has to undergo a change from a longer form of LLM to a shorter one when appearing immediately before another tone, in order to comply with the tim-

ing requirements stipulated in the Principle. Interestingly, the implementation rule involved in 3TS(A) (i.e. a rule that modifies Tone 3 to satisfy the Principle) is exactly the same rule as the one seen in the case of 4TS; that is, the Tone Truncation Rule (henceforth TTR) in (6.7). The 3TS (A) derivation via the TTR is shown below:

(6.11).



As clearly shown here, the TTR, when applied on a Tone 3 followed by another tone, yields the exact grammatical result LL.

Notice how the present theory unifies the two sandhi processes, 4TS and 3TS(A), in a principled way. These sandhi rules now become one and the same shortening process operating under the influence of the Mandarin-specific Tone Reduction Principle. The general nature of the present analysis stands out in contrast to previous analyses, all of which treat these two sandhi processes as being accidental and unrelated.

As in the case of 4TS, naturalness is again shown here in the present analysis for 3TS(A). The toneme apocope is naturally performed by the TTR, as naturally as a final segment is deleted in a segmental analysis. It should now be more clear that without treating the longer form of Tone 3 as the underlying one, the process of 3TS(A) cannot be as naturally captured, neither can the connection between 4TS and 3TS(A).

As in the case of 4TS, the present analysis explains rather than simply describes 3TS(A). With the postulation of the Principle, the reason for 3TS(A) to happen is just that it is followed by another tone; naturally, due to the presence of a following tone, the first of the two juxtaposed Tone 3s must drop one of its timing units.

These characteristics of my analysis can be shown further if it is compared with some pertinent previous analyses. First of all, none of the previous studies brings out the insight that it is simply because of the presence of a following tone that the affected Tone 3 undergoes sandhi. Chao, by treating (213) as underlying, could potentially reveal this insight; however, with a tone-replacement treatment, such an insight is concealed.

The problem with Woo and Yip begins with their treatment of the shorter LL (21) rather than the longer LLM/LLH as the underlying form. With such a treatment, they have to employ an ad hoc rule (given below) to add a toneme to Tone 3 when it occurs in a final position:

(6.12). Woo: LL → LLM / ___ (pause) (cf. 2.15)
or, Yip: LL → LLH / ___ (pause)

The ad hoc nature of the rule is shown first of all by the fact that it is too case-specific, exclusively devised for only a single sandhi process involving only Tone 3, without any independent motivation. Secondly, by adopting this rule, Woo's and Yip's analyses become quite unsymmetrical: they indicate that Tone 3 alone from among all the Mandarin tones, receives an extra toneme in an utterance final environment. It seems that for their analyses to be adequate, they would need to explain why the other Mandarin tones do not acquire such a tail in utterance final position. Notice that by treating the shorter form as the underlying form, Woo and Yip are in effect making the incorrect claim that there is no SANDHI process here in the sense that a tone is affected by the presence of another tone in its periphery. And by arbitrarily attaching a H (or M) to the form LL, their analyses imply that the alternation between LLH (213) and LL (21) is accidental, and thus fail totally to capture the reason behind the alternation.

Of the two, Yip shows yet another inconsistency with the treatment of the shorter form of Tone 3 as underlying: in utterance final position, Tone 4 has the

value of (51) while Tone 3 has the value of (213), yet Yip has only a rule to create (213) but not one to do similar justice to (51). The inconsistency seems to have resulted from the failure of her theory to provide a theory-internal answer to the longer forms of the Mandarin tones.

6.2.3 The Third-Tone Sandhi: B (3TS(B))

Besides (21), Tone 3 has yet another alternant (35), considered identical to Tone 2 (35) in value. The alternation is seen by Chao as involving the following process (also see 6.1c):

(6.13). The Third-Tone Sandhi: B (3TS(B))

213 → 35 / ___ 213

That is, 3TS(B) involves the derivation of (35) on a Tone 3 when this Tone 3 occurs before another Tone 3. Examples of this sandhi process are given below:

(6.14).a. nǐ(213) hǎo(213) →
hí(35) hǎo(213) "hello"

b. lǎo(213) hǔ(213) →
láo(35) hǔ(213) "tiger"

c. mǎi(213) jiǔ(213) →
mái(35) jiǔ(213) "buy alcoholic drinks"

d. wǔ(213) zhǒng(213) →
wú(35) zhǒng(213) "five kinds"

The nature of this sandhi process can again be captured by the Principle. But before the presentation of my analysis for 3TS(B), let us first have a look at how it is handled in previous analyses. I will here focus on Yip specifically (1980a, 1989a,b). This is because that the problems with Chao's tone-replacement model will become fairly self-evident as the discussion of the process proceeds. And for a discussion of the problems with Woo's analysis, please refer to § 2.5.

As mentioned earlier, Yip's representation of Mandarin tones consists of two tiers, Register Tier and Tonal Tier. Representing Tone 3 as [-upper, LL], Yip (1980a) gives the following rule for 3TS(B) (p.183):

(6.15). [-upper] → [+upper] / ___ [-upper]
L L → L H → L L

According to Yip, the above rule can be separated into the following two rules, each working on a separate tier:

(6.16). a. [-upper] → [+upper] / ___ [-upper]

b. L → H / [+upper] L ___

Let me first discuss the rule in (a). This rule says that when one [-upper] precedes another [-upper], the first one changes into [+upper]. According to Yip, this rule is dissimilatory, and it is a rule whose operation is triggered by the OCP. If Yip is correct here, she then needs to explain why the OCP fails to work on sequences of tones such as the following:

(6.17). a. Tone 1 + Tone 2
[+upper] [+upper]
H H L H

b. Tone 2 + Tone 1
[+upper] [+upper]
L H H H

c. Tone 1 + Tone 4
[+upper] [+upper]
H H L L

d. Tone 4 + Tone 1
[+upper] [+upper]
H L H H

e. Tone 2 + Tone 4
[+upper] [+upper]
L H H L

f. Tone 4 + Tone 2
[+upper] [+upper]
H L L H

g. Tone 1 + Tone 1
[+upper] [+upper]
H H H H

In each of the above seven cases, a sequence of two tones with an identical Register value of [+upper] is seen. Yet, in none has a dissimilation process occurred. If the OCP should work on a sequence of two lower Register values, it would seem odd if it fails to work on a row of two higher Register values. One would probably need to explain what is special about the lower Register (i.e. [-upper]) that causes it to have an exclusive sensitivity to the OCP.

If one argues that the reason by which the OCP fails to intervene in the cases (a)-(f) is that the two tones involved in each case are all non-identical at the tonemic level, one would still need to explain the (g) case, in which both the Register and the Tonal values are identical. It seems that the only explanation one can give here, in view of (g), is that the OCP only works on tones of lower Register and not on higher ones. This explanation, however, would sound unconvincing. No other evidence is provided by Yip to show that the OCP works in such a selective way.

A similar problem is found in the other half of the 3TS(B) rule proposed by Yip. This second rule (6.16b) says specifically that when two tauto-tonic tonemes dominated by the same higher Register node are identical LL tonemes, the second dissimilates and changes into H. According to Yip, the alternation is again caused by the OCP. The first problem with this rule comes from the OCP working on a sequence of two identical tonemes which belong to the same tone. As is commonly recognized, one major difference between Chinese tones and tones in the African languages is that the former, though characterizable in terms of a sequence of tonemes, often exhibit a holistic nature not found in the latter in which tauto-syllabic tonemes may behave like individual tones.

The point is that a sequence of identical tauto-tonic tonemes in Mandarin (these tauto-tonic tonemes are always tauto-morphemic as well) actually form one

continuous articulatory unit unaffected or unbreakable by principles such as the OCP. Tone 1, for instance, though represented by more than one toneme in a non-linear treatment, is nonetheless just one continuous level tone. In spite of its representation, it should not be forgotten that the purpose of a representation with multi-tonemes of identical value is to represent duration.

A parallel situation to such a toneme cluster can be found in what have been labelled as "true geminates" (Hayes 1986, Goldsmith 1990). According to Goldsmith, true geminates are those "that are internal to a single morpheme (tauto-morphemic), [while those] formed across a morpheme boundary are only apparent geminates (p.81)." I believe that the former is also true of the toneme clusters, LL or HH, etc., under consideration. That is, these toneme clusters are also true toneme geminates, as they are tauto-morphemic as well.

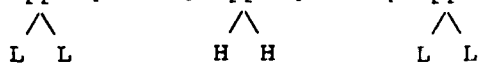
According to Goldsmith and several other studies (e.g. Hayes 1986, Itô 1988), true geminates have the qualities of integrity and inalterability. True geminates, for example, are found distributed in positions where a cluster of two distinct consonants are disallowed. They are also found to be resistant to vowel epenthesis rules even if they meet the structural descriptions of these rules. In a Semitic language called Trigrinya (Hayes 1986), for example, the tauto-morphemic geminate /kk/ fails to be altered by a rule which otherwise alters a /k/ to /x/. Above all, true geminates are found not subject to the influence of the OCP. All these facts suggest that in the case of true geminates, the use of two identical symbols is simply meant to represent a durationally longer consonant or vowel. Hence, just as a true geminate is insensitive to the OCP, these toneme geminates (or toneme "trios" with three identical tonemes) should be equally unaffectable.

Arguably, at least in a tonal system like the one in Mandarin, the intervention of the OCP happens only at the "breaks" of a two identical-toneme sequence found between tones (or syllables).

One may already have noticed that this rule of Yip (i.e. 6.16b) is in any event too specific. It handles a specific case in a specific environment. There is no independent evidence to support it. In particular, the specific case is a L (not H or anything else) toneme dominated by specifically a higher Register (excluding a lower Register), and the specific environment in which this L toneme changes is one immediately after another L (not H or anything else) dominated by the same Register node.

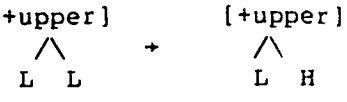
The most serious problem for this rule lies in the representation of Tone 1 and Tone 3 within the same framework. First of all, if the OCP indeed works on a sequence of two L tonemes dominated by a higher Register node, one should expect it to work on a sequence of the same two L tonemes dominated by a lower register node. Or, one should expect it to work on a sequence of two H tonemes dominated by the same higher register node. In the following, all these three situations are given formally in Yip's analysis:

(6.18). a. [+upper] b. [+upper] c. [-upper]



One common feature among the three cases is that all are level tones represented by two identical tonemes. Yip's rule (6.16b) requires the alternation of specifically (a) case, as shown below,

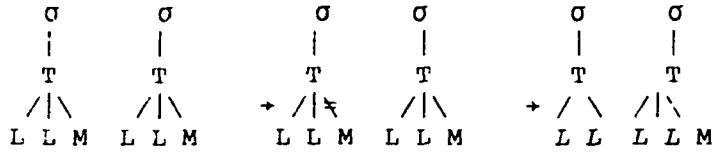
(6.19).



while saying nothing with regards to (b) and (c). However, as it turns out, these latter cases do occur, and in fact, they are the analyses for Tone 1 and Tone 3 respectively in Yip. It seems that one must explain why (b) and (c) occur at all if one maintains (6.16b). It is not clear to me if such an explanation can be found. Without an adequate explanation for (b) and (c), the rule in (6.16b) cannot be accepted.

Now let us turn to my analysis for 3TS(B). As it turns out, 3TS(B) is also explicable by virtue of the Principle stated in (6.4): Tone 3, represented as LLM, must undergo reduction; it has to be "tailored" to two tonemes, just as the affected tones do in the case of 3TS(A) and 4TS. And, just as in the latter two cases, the "tailoring" is conducted by the TTR, and the derivation involved is as shown below:

(6.20).



As is expected, 3TS(B) involves more than TTR. This should not be surprising since there would not be two 3TS processes if only this TTR were involved. The question is: what is the difference between this 3TS(B) process and the 3TS(A) process seen in the last section? My answer is that while in the latter, the application of TTR yields a grammatical result, the result after TTR in the former, where identical elements are found at the boundary between the two juxtaposed Tone 3s, violates an important, widely observed constraining condition: the Obligatory Contour Principle.[10]

But, how exactly does the effect of the OCP manifest itself here, since what is seen here is a sequence of four identical tones which allows for a number of possible explanations? As I mentioned earlier, there are reasons to believe that in Mandarin, the OCP only works at the tonal boundaries and not between the tautomorphemic toneme geminates. Given this position as a condition, there are four possibilities left. One, the OCP does not allow a sequence of LL LL, namely, a long low tone followed by another long low tone. Two, the OCP does not allow LL L, namely, a long low tone followed by a short low tone. Three, it does not like L LL - a short tone followed by a long low tone; or four, the OCP does not permit L L, namely, a short L tone followed by anther short L tone.

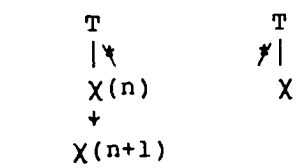
Among the four possibilities, the last one stands out as the most generalized one. That is, what is true for this combination of tonemes is also true for all the other three. Thus, if the last one can be established as the condition here, it would yield a more generalized theory. In fact, it is indeed this most generalized situation which constitutes an OCP violation in Mandarin. This OCP-violation situation can be expressed in terms of the following Mandarin-specific Well-Formedness Condition (WFC)[11]:

(6.21). *

where χ = H, M, or L

What this WFC says is that two adjacent tonemes, whether H, M, or L, are not permitted if they are dominated by different tones (i.e. T nodes). Thus, due to this WFC, the output of the TTR on a sequence of two Tone 3s has to be repaired, and the repair is evidently done to the first of the two non-tauto-tonic tonemes by the following rule,

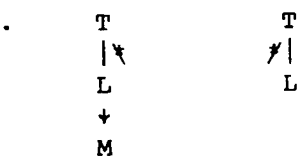
(6.22).



where χ = H, M, or L,
and n = pitch level
thus, for L, n=1
M, n=2
H, n=3

and in the following manner:

(6.23).



That is, one of the offending tonemes, the left one in this case, is raised by one degree in pitch level. The dissimilatory result that is achieved on the two Tone 3s in question is now LM LLM.

Here, one may have noticed what seems to be a problem in the analysis proposed so far: after the dissimilation, a rising tone in the shape of LM is obtained rather than one in the shape of MH. Should one then decide that the present theory fails to derive the correct surface form, and should be discarded accordingly? The answer is no.

Several considerations have in fact led to the assumption that after the dissimilation, the resultant rising tone is a lower LM rather than a higher MH, and that this result of LM is then merged (or defaulted to) with MH (i.e. the value of Tone 2 in the language). For one thing, MH, being the value of Tone 2, is an already existent value of the grammar. It is possibly the only value a Mandarin speaker knows how to produce in pronouncing a rising tone. Because of this and possibly because of a linguistic tendency toward economy in the grammar, any rising tone in Mandarin, no matter what value it has and through what derivation process it is derived, will be elevated to merge with this higher rising tone MH.

Unfortunately, there does not exist in Mandarin another tone sandhi process which also derives a rising tone to test the proposed merging process. However, supporting evidence can be found in the Tianjin-based dialect of Mandarin. According to Hung (1987), there are four tones in Tianjin:

(6.24). Tone A: 21
Tone B: 45
Tone C: 213
Tone D: 53

Operating on these four tones are four tone sandhi processes, two of which happen to produce rising tones:

(6.25). 213 → 45 / ___ 213
53 → 45 / ___ 21

No matter what the base tone is, (213) or (53), and no matter what the sandhi environment is, the rising tone that is derived from each case is none other than the lexical rising tone (i.e. Tone B) of this dialect. What this observation suggests is, as I just proposed, that there is in Mandarin a tendency, if not a general rule, for the rising tones to merge together no matter what their original shapes are. From this, one probably can also expect derived tones of other tonal categories than the rising tone to merge with an inherent lexical tone of the same contour in the same language.

Strong evidence for this merging is also found in an acoustic study of Mandarin tones by Shen (1990b). Results from Shen's instrumental study indicate that the derived rising tone is a lower rising tone before it merges with the inherent rising tone (i.e. Tone 2). In this study, which focuses on the co-articulation effects among Mandarin tones, Shen makes the following interesting observation:

> "Although after [3TS(B)], the tonal contour of the first Tone 3 is similar to Tone 2 in that it becomes a rising tone, the second Tone 3 is higher when following an inherently rising tone than a rising tone generated by [3TS(B)]." (p. 285)

How then can one explain the observed difference between the varied pitch values of the Tone 3 after a real Tone 2 and the Tone 3 after a derived Tone 2? Shen suggests that "Tone 3 after tone sandhi does not rise as high as Tone 2 (p.285)." If Shen's explanation is valid, there is one more reason to assume the earlier existence of a lower rising tone LM that becomes identical to the inherent rising tone only afterwards. What I mean is a rule that says the following:[12]

(6.26). Tones of the same contour tend to merge together.

This way, one can give the explanation that the second Tone 3 is lower after a derived rising tone than after an inherently rising tone because the derived rising tone, though ultimately merging with the value of the inherent rising tone, is nevertheless not quite the same phonetically as the inherent rising tone. When a derived rising tone is a lower rising tone in the first place, it is only natural for it to exert a lower co-articulation effect on the following third tone than that exerted by the inherent rising tone. By this, I assume, of course, that the merging process is a lower phonetic process than the co-articulation process, which means that the former is ordered after (or may be simultaneous to) the co-articulation process. Thus, if Shen's explanation is valid, as I believe it is, it should pose a problem for the previous analyses which equate absolutely the derived Tone 2 with the inherent Tone 2. By saying simply that "Tone 3 becomes Tone 2", Shen's observation is not explained.

There still seem to be counter-examples to the WFC encoded in (6.21). These examples involve Tone 1, Tone 2, and Tone 4, the tones with either a H onset (Tone 1 and Tone 4) or a H offset (Tone 1 and Tone 2). The problem is that when these tones are strung together in some fashion, the structural description expressed in (6.21) is met, signalling an OCP violation.

(6.27). Tone 1 + Tone 1
HH HHH
Tone 1 + Tone 4
HH HML
Tone 2 + Tone 1
MH HHH
Tone 2 + Tone 4
MH HML

The question that must to be answered is why the OCP seems to have failed to intervene in these four cases. My answer to this question is that the OCP repair rule in (6.22) does not cover these cases. Recall that this repair rule raises the tonal level of the offending toneme by one degree. However, the offending

tonemes in all of these above cases are already maximally high tonemes in the present theory, thus leaving no room to be increased further in pitch. In other words, it is not that the repair rule has not applied and therefore the OCP has failed to intervene, but that the repair rule applies only vacuously on these four cases, leaving them unchanged.

Notice that the same features found in the present analyses of 4TS and 3TS(B) are also true here. For one thing, the two 3TS processes have never been associated with each other in the past, but they have been treated as two isolated processes. Within the present theory, both (as well as 4TS) undergo the same rule (i.e. TTR). The difference is just that one (i.e. 3TS(A)) comes out of the rule without violating any language-universal or language-specific WFCs, whereas the other (i.e. 3TS(B)) coming out of the rule with an OCP violation that has to be fixed.[13]

6.2.4 The Second-Tone Sandhi (2TS)

Besides Tone 3 and Tone 4, the Second Mandarin Tone has also been found to undergo change in its tonal value in a certain environment. Chao (1968) has described the process as follows:

> "A tone sandhi of minor importance has to do with the change of the 2nd to a 1st in three-syllable groups." (p.28)

Cheng (1973) has made similar descriptions:

> "In fast conversational speech, a second tone becomes first when preceded by first or second tone and followed by any tone other than the neutral tone." (p.44)

The process can be roughly delineated as follows (also see 6.1d):

(6.28). The Second-Tone Sandhi (2TS)

35 → 55 / {35, 55} ___ {55, 35, 213, 51}

Note: This rule applies in fast speech only.

Examples of 2TS are given below (adapted from Chao, p. 28):

(6.29)

a. xī yáng shēn → xī *yāng* shēn
   occidental ginseng "(occidental) ginseng"

b. sān nián jí → sān *niān* jí
   three year grade "grade three"

c. cōng yóu bǐng → cōng *yōu* bǐng
   onion oil pancake "a kind of pancake"

d. dōng hé yànr → dōng *hē* yànr
   east river bank "east riverside"

e. fēn shuí líng → fēn *shuī* líng
   separate water shed "watershed"

f. shuí néng fēi? → shuí *nēng* fēi?
   who can fly "who can fly?"

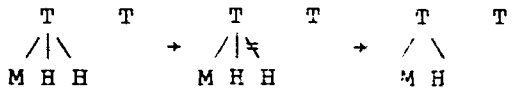
g. hái méi wán → hái *mēi* wán
   still not finish "not yet finished"

h. yóu zhá huì → yóu *zhā* huì
   fried dough-nut "fried dough-nut"

Unlike the other three sandhi processes, this 2TS, as far as I know, has not received any formal treatment. The lack of a treatment for this sandhi process is probably attributable to a lack of an understanding of the process involved. After all, nothing interesting can be said about it if it is considered merely a tone-replacement process which substitutes Tone 1 for Tone 2. Besides, there is clearly no dissimilation involved, as there is in the cases of 3TS(B). Nor is there an obvious loss (or addition) of a "tail" as seen in Tone 3 in 3TS(A). In short, no obvious relation can be built between the two alternants of Tone 1 and Tone 2.

Within the present analysis, this 2TS process becomes no more mystical than the other three tone sandhi processes. Just as in the other three processes, a tone that is followed has to undergo the same reduction of a toneme. Let me address the affected medial Tone 2 first. This tone, being followed, has to undergo the TTR as seen in the previous three cases. The derivation is as shown below:

(6.30).



Through TTR, a resulting shortened rising tone (MH) is acquired. Apparently, more than TTR is responsible for the derivation of the correct surface result. This should not be surprising, as the affected Tone 2 is in a context surrounded on both sides, while the Principle as stated in (6.4) only covers the following environment of a tone without any mentioning of its preceding environment. That principle, as has been shown, works well for 4TS, 3TS(A) and 3TS(B) because in these cases, only two tones (or syllables) are involved, and the change of the first tone requires merely the presence of the second tone and no more. Compared with 4TS, 3TS(A) and 3TS(B), 2TS has one further condition on it: being at the same time preceded by another tone.

Logically speaking, if the presence of a following tone should have a shortening effect on its preceding tone (within the same foot, of course; cf. § 1.3), it should not be surprising to see similar effect from a preceding tone on a following tone (iff of course all these three tones also simultaneously belong to the same foot). Hence, I propose to extend the Principle as follows to include an account (Clause B) for the effect coming from a preceding tone:

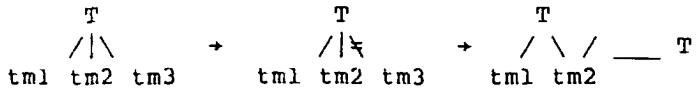
(6.31). Tone Reduction Principle

> Clause A: (= 6.4)
> In normal speech, reduce a tone by one toneme iff it is immediately followed by another tone within the same prosodic foot.
>
> Clause B:
> In fast speech, reduce a tone by one toneme iff it is immediately preceded by another tone, and at the same time immediately followed by another tone within the same prosodic foot.

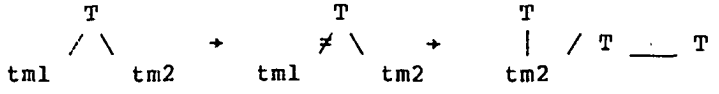
Furthermore, just as Clause A of the Principle has a rule (i.e. TTR) to implement it, so should Clause B have a similar implementing rule. This rule, more or

less symmetrical to TTR, may be referred to as the Rightward TTR (see below in 6.30b) as opposed to the TTR seen earlier which I may now refer to as the Leftward TTR (repeated here as 6.30a):

(6.32).a. Leftward TTR (L-TTR)



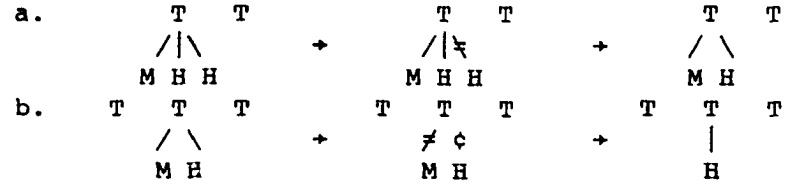
b. Rightward TTR (R-TTR)



Notice that R-TTR is ordered after L-TTR. This means that only if the former is applied will the latter be applied. This is because that only if a tone is followed at the same time will its preceding tone have an effect on it. In a two syllable word such as *yǔyán* "language", the second syllable (i.e. the final one) is not affected even if it is preceded.

Now if both rules in (6.32) are applied to Tone 2 in the 2TS environment in the order just discussed, an output of a high level tone is achieved:

(6.33).



As I will argue below, this result of a H tone is precisely the correct outcome of 2TS.

Recall that in the traditional description of 2TS, it is said that the result of 2TS is a Tone 1 (see the beginning of this section for Chao's and Cheng's remarks). Such a description implies that the tone derived through 2TS is identical to the underlying Tone 1 in every respect. It is identical not only in pitch value (5 on

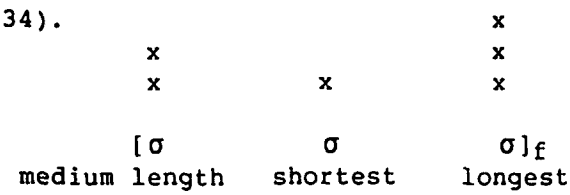
Chao's scale) but also in length to an underlying Tone 1. However, as I mentioned earlier, studies have shown that the medial tone of a three-tone foot (i.e. a super-foot; cf. § 1.3) is considerably shorter than tones in the other two positions. It is in fact the shortest among the three, and therefore, much shorter than an underlying Tone 1 (in final position).

For instance, Chao (1968) has made the following observation when talking about the stress pattern of multi-syllabic units:

> "Sequences of normally stressed syllables without intermediate pause, whether in a phrase or in a compound word, are not all of the same degree of phonetic stress, the last being the strongest, the first next, and the intermediate being least stressed." (p. 35).

In terms of the physical realization of a stress, Chao remarks in the same page that "stress in Chinese is primarily an enlargement in ... duration (p. 35)." Thus, it is clear from Chao's observation that in a three-syllable foot, the last syllable is the longest, the first next, and the medial shortest, as shown below:[4]

(6.34).



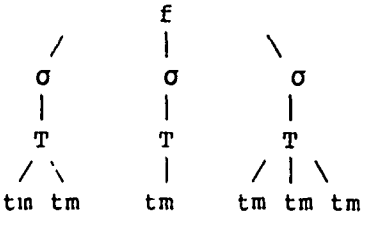
Chao's observation of such a stress pattern is supported by Yan and Lin (1988). In Yan and Lin, sonagraphic analysis of utterances of three-syllable units made by native Mandarin speakers confirms that under normal circumstances, a three-syllable unit has a stress pattern such as this: medium, light, and heavy. Shen's (1990b) research provides additional support for such a pattern on three-syllable units. In her study of co-articulation effects among tones, Shen has her subjects

utter a three-syliable string, /pa pa pa/, in various tonal combinations, and records the utterances. Before the recording, the subjects are specifically "instructed to stress the three syllables evenly and not to pause between them (p. 283)." In spite of the instruction, Shen finds that "due to articulatory habits, a number of tokens were produced following the most frequent stress pattern of trisyllabic words/phrases, that is, with the last syllable most stressed and the second syllable least stressed (p. 283)."

If Chao is correct, as I believe he is, in maintaining that stress positively correlates to duration, the findings of these studies by Chao, Yan and Lin, and Shen concerning the stress/duration patterns fall out from the present analysis of the 2TS. As I have already shown, the medial tone came out in the present analysis as the shortest tone of the three.

Notice that the present analysis does not just take care of the medial tone, but it also correctly modifies the first of the three tones. It is clear that by virtue of the Principle, the first of the three (with three underlying tonemes) has to be modified as well, just as the affected tones do in 3TS(A)-(B) and 4TS. The result is, as expected, a two-toneme tone. Now let me sum up the result of the present analysis for a three-syllable string as follows:

(6.35).



Now, compare (6.34) and (6.35), one can see that the present treatment yields the correct result of duration.

Compared with previous treatments of 2TS, the present analysis is explicit. Rather than coarsely saying that the affected Tone 2 in 2TS becomes Tone 1, the

present analysis reveals that the result is after all not quite Tone 1 in its underlying form: it is identical to an underlying Tone 1 only in pitch value, but not in duration. Rather than saying in a crude way that Tone 2 becomes Tone 1 in certain context, our analysis reveals why Tone 2 becomes Tone 1 in that context, how exactly this derived Tone 1 is related to its original Tone 2. Under this analysis, the derived high level tone, or Tone 1 in traditional terms, does not just come from nowhere, as the previous treatment would indicate. It is but part of the original Tone 2, or "left over" from it, after some natural truncation processes have taken place. The problem with previous treatments is that they do not have an explicit formalism to account for duration so that they have to describe the result of 2TS as equatable with an underlying Tone 1. Another feature of the present analysis of 2TS is, once again, its generality. It relates, in a principled way 2TS to the previously-discussed tone sandhi processes in Mandarin, so that a unified account for all these process becomes now possible.

To conclude this section, I would like to point out two things. One, within the present analysis, not just Tone 2, but every Mandarin tone, is modified in the medial position. Two, not only do Tone 1 and Tone 2, but all the four phonemic tones, have a shortening effect on the following medial tone. Let me address the former first. The reason that only Tone 2 has been observed to be modified to a level tone is simply because that modification is intuitively more observable. Recall that Tone 1 (HH) and Tone 3 (LL) are both level tones before the modification (by 31.b). Therefore, a change which simply transfers these level tones to shorter level tones is not intuitively observable, nor is it observable or describable by an analysis which does not encode duration.

It is not clear to me why Tone 4 has not been observed to be modified in that situation, although clear evidence suggests that Tone 4 is also flattened to a level tone in the medial position in fast speech. The following are a few examples:

(6.36).

a. fāng kuài zì → fāng *kuai* zì
square piece character "square/Chinese character"

b. fen jiè xiàn → fēn *jie* xiàn
separate border line "border line"

c. lián yì huì → lián *yi* huì
connect friendship association "friendship association"

d. míng xìn piàn → míng *xin* piàn
open letter card "postcard"

e. chá huà huì → chá *hua* huì
tea talk meeting "tea party"

f. lǎo dà gē → lǎo *da* gē
old big brother "big brother"

g. kǒu qì dà → kǒu *qi* dǎ
mouth air big "talk big"

h. kě kào xìng → kě *kao* xìng
may rely quality "reliability"

i. lǐng shì guǎn → lǐng *shi* guǎn
consul office "consulate"

j. rǔ zhì pǐn → rǔ *zhi pǐn*
milk make product "dairy product"

k. lěng bù fáng → lěng *bu* fáng
cold not protect "all of a sudden"

l. mài kè fēng → mài *ke* fēng
microphone "microphone"

m. dì zhì xué → dì *zhi* xué
land quality study "geology"

n. fù shì zhǎng → fù *shi* zhǎng
vice city head "vice mayor"

o. kàn bù qǐ → kàn *bu* qǐ
see not rise "look down upon"

p. fù zuò yòng → fù *zuo* yòng
side effect "side effect"

q. qù qù qù → qù *qu* qù
go go go "go away!"

In all of the above cases, the medial Tone 4 is pronounced, in fast speech, as a contourless tone (signalled by not being marked by any diacritics), just as Tone 2 is. These data also show that the conditioning initial tone can not only be Tone 1 and Tone 2, but Tone 3 and Tone 4 as well. The reason that only the former two tones have been observed to affect the medial Tone 2 is probably due to the fact that the end points of the latter two tones are made up of lower tonemes which may have lower co-articulation effect on the following Tone 2 so that Tone 2 does not exactly surface phonetically with a perfect high level shape.[15]

6.2.5 The Neutral-Tone Sandhi (0TS)

In § 3.1, it was argued that Yip's (1980a) analysis of Tone 0 is inadequate. This section provides evidence to further confirm that conclusion. In particular, I will show that Yip's analysis fails to capture an important generalization concerning Tone 0. In fact, not only Yip, but all the previous accounts of Tone 0 suffer the same flaw, which is due evidently to the misleading pioneer description of Tone 0 given by Chao (1968). According to Chao, Tone 0 does not have a phonemic tone shape (or pitch value) of its own; rather, it derives its pitch value from tones on its periphery. When following other tones, it has the pitch value that varies with the pitch values of these preceding tones. Chao gives the following description of the varied tonal values of Tone 0.

(6.37).

		Tone 0
Tone 1	55	2
Tone 2	35	3
Tone 3	213	4
Tone 4	51	1

It seems that Chao's pioneer work has led many subsequent analyses of Tone 0 to model on this description and to try hard to derive the various pitch values of Tone 0 through phonological processes. By so doing, these analyses have missed an important and otherwise fairly obvious generalization about these absolute pitch

values. This generalization, as I will soon demonstrate, can only be brought to light if Tone 0 is treated as having a specific underlying value of its own.

It has been noted in several studies that Tone 0 is characteristically low after every other tone except Tone 3. Cheng (1973), for instance, maintains that

> "the refined acoustic details [about the neutral tone] perhaps do not necessarily represent the native speaker's knowledge. The speaker's aim perhaps is to produce the neutral tone low after first, second, and fourth tones and higher after third tone." (Cheng, p. 56)

As a matter of fact, the same idea is also implied, though in a rather intuitive manner, in Chao when he gives the advice that "for practical purposes, it is sufficient to remember the neutral tone as being high after a half third tone[16] and (relatively) low after the other tones (p. 36)."

The special behavior of Tone 0 after Tone 3 is expressed not only in such intuitive remarks as given above,. it is also corroborated by findings from empirical research. In their acoustic study, Dreher and Lee (1966), for example, find that it is only after Tone 3 that Tone 0 rises, while after any other lexical tones, it falls. Their findings are given below:

(6.38).

	Tone 0
after Tone 1 (55):	41
after Tone 2 (35):	31
after Tone 3 (21):	23
after Tone 4 (51):	21

Dreher and Lee's findings are further confirmed by an instrumental study by Gao (1980) who also finds Tone 0 falls after Tone 1, 2 and 4 but rises after Tone 3. In still another study, Dow (1972) finds that Tone 0 is uniformly low after Tone 1, 2, and 4, but relatively high after Tone 3. His illustration (p. 113) is provided below:

(6.39).

	T1+T0	T2+T0	T3+T0	T4+T0
200-				
180-				
160-			*	
140-				
120-				
100-	*	*		*

On a scale from 100 to 200, Dow marks all the neutral tones after Tones 1, 2 and 4 at the 100 level[17], with the single exception of Tone 3. After Tone 3, it is at the level of 160. In fact, the findings in Dreher and Lee are identical to Dow if the onset values of Tone 0 in Dreher and Lee are ignored. There, at the offsets, one sees a pitch value of 1 in the environments after T1, T2 and T4, but a pitch value of 3 after T3.

These above-mentioned studies confirm the intuition expressed in Cheng and Chao that T0 is high after the third and uniformally low after the other three phonemic tones. What, then, causes Tone 0 to act thus differently? Or, rather, what is it in Tone 3 that causes the following neutral tone to act differently? I argue that the answer to this question lies in the following assumption of the representation of the neutral tone.

(6.40). The Representation of Tone 0

Given this representation[18], the mysterious behavior of Tone 0 after the third becomes easily explicable. Now let us have a look at the tonemic representation of Tone 0 following other tones in the underlying representation:

(6.41).

		Tone 0
Tone 1	HHH	L
Tone 2	MHH	L
Tone 3	LLH	L
Tone 4	HML	L

Recall that the Principle in the present analysis cuts the tail of a tone when it appears before another tone. This means that (6.41) really looks like the following, where the last tonemes of the preceding tones are deleted (by TTR, the rule that implements the Principle):

(6.42).		Tone 0
a. Tone 1	HH	L
b. Tone 2	MH	L
c. Tone 3	LL	L
d. Tone 4	HM	L

It is seen in the above a problem with the Well-Formedness Condition independently motivated previously in the discussion of the 3TS(B). To refresh our memory, let me repeat the WFC here:

(6.43). * T T
|\ /|
X X

where X = H, M, or L

The problem seen in (6.42c) shows up clearly when evaluated against this WFC. The end point of Tone 3 is identical to Tone 0 in pitch, in violation of the WFC. Thus, given the WFC, a dissimilation process has to be invoked to repair (6.42c), just as in 3TS(B). The repair is, however, done in a slightly different manner, as shown below in (a):

(6.44).a. T T
|\ /|\
X X
+
X+1

b. T T
|\ /|
X X
+
X+1

where X = H, M, or L

Now compare (a) with (b), the latter being the OCP-violation repair rule involved in 3TS(B). One can see that there is only a minor difference existing

between the two. While in the case of 3TS(B), the repair (b) is done to the left of the two offending tonemes, in the case of Tone 0, the repair (a) is done to the right one. Unfortunately, I have no particularly strong explanation for this difference, except the speculation that in the case of Tone-0 modification, the toneme being modified stands alone on its own without being associated with any other toneme under the same tone node, and is, therefore, phonetically less stable.

Notice that the two rules in the above consist of the two and only two possible repair alternatives. Formally speaking, the distinction between these two rules is that (a) is more specific, while (b) is an elsewhere rule.

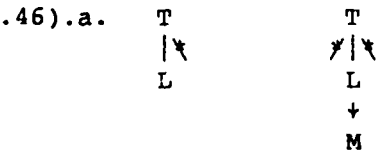
Now let us see the derivation on Tone 0 by the rule given in (6.46a). With this rule, Tone 0 after Tone 3 is raised from L to M, as expected,

(6.45).

		Tone 0	
a. Tone 1	HH	L	
b. Tone 2	MH	L	
c. Tone 3	LL	M	(L→M)
d. Tone 4	HM	L	

in the following manner:

(6.46).a.



If the present analysis of the basic values of Tone 0 is correct, there is one more tone-sandhi phenomenon in Mandarin. It may perhaps be called the Neutral-Tone Sandhi process and delineated roughly in Chao's notation as follows:

(6.47). The Neutral-Tone Sandhi (0TS)

(1) -> (3) / (21) ___

Now that the value of Tone 0 after Tone 3 is accounted for, let us see if the analysis also holds water when it comes to the surface values of Tone 0 after the other tones. There, it is found that only when the present analysis is adopted is it

possible to give a principled explanation for these other bewildering, varied pitch values of Tone 0.

Recall that according to some instrumental studies, Tone 0 rises in pitch after Tone 3 but falls after the other tones. Before, when Tone 0 was not given an underlying representation as it is given here, this fact was totally inexplicable. This is probably why previous studies have merely described the variance, but not provided any explanation for it. It becomes fairly apparent under the present analysis that the rise and fall of Tone 0 is due to low level co-articulation (or assimilation), which occurs after 0TS.

In the case of Tone 0 after Tone 1 (6.42a), the toneme before Tone 0 is a H. From this H to the L of the Tone 0, a natural phonetic fall is inevitable. The same is true in the case of Tone 2 (6.42b), the offset of which is also a H. Shen's (1990b) remark lends support to the existence of such an co-articulation effect:

> "Tones 1,2,4 and 0 have their highest overall tonal value when following the two high-offset tones, Tones 1 and 2." (p. 284).

In the case of Tone 0 after Tone 4 (6.42d), the toneme before Tone 0 is a M. From this M to the L of the Tone 0, a natural fall is again inevitable. In addition, given the fact that the end point of Tone 4 is lower than the end point of Tone 1 or Tone 2, it is only expected that the absolute value of Tone 0 after Tone 4 is lower than those after Tone 1 and Tone 2. One puzzling observation made in several instrumental studies (e.g. Shih 1987) is that Tone 4 falls lower than normal before Tone 0. This observation is again only explicable if Tone 0 is assumed to contain in the underlying representation a L toneme. Tone 4 falls more than usual is because of the anticipatory co-articulation effect exerted on it from the low toneme of the Tone 0 that follows it. Finally, the rise in Tone 0 after Tone 3 becomes, by now,

fairly trivial: it is but the result of co-articulation between the L toneme of Tone 3 and the raised toneme of Tone 0, that is, M.

Conversely, the present analysis is able to predict the influence of the co-articulation effect. Due to the co-articulation effect, it is only expected that the starting point of the Tone 0 should vary in pitch value, in view of the varied pitch value of the preceding toneme. In particular, the starting point of the Tone 0 should be higher after Tone 1 and Tone 2 but lower after Tone 4. These are exactly the measured results in the instrumental studies by Dreher and Lee (1966), and Gao (1980), as well as from Chao's intuitive observation.

So far, I have demonstrated that the much observed data concerning Tone 0 become tractable only when one assumes that Tone 0 is underlyingly a low tone. Now, I would like to try to give an articulatory explanation for the low nature of Tone 0. First of all, let me try to describe what happens to the articulators when the neutral tone is uttered.

To a native speaker who has no knowledge of linguistics, Tone 0 is the same neutral tone with the same pitch level INTENDED no matter where it appears. In a sense, one may say that Tone 0 is not really an intended tone, and that in the production of it, a particular pitch level is not targeted at. Although it shows up with a low pitch that can be described as a low tone, the origin of the pitch is nothing like the intended pitch levels in other tones. This seems a feature that characteristically distinguishes Tone 0 from the other four lexical tones.

Physiologically, the articulation of such an unintended pitch level is characterized by the laxing of the laryngeal and/or super-laryngeal muscles caused by a lack of stress in the environment where the neutral tone appears. A comparable situation can in fact be found in the well-attested process of schwa reduction typical in many of the world's languages. In English, for example, a vowel, no matter

what shape it has, tends to be reduced to a schwa in unstressed or weakly stressed position. The nature of the reduction is again attributable to the laxness of the pertinent muscles from lack of stress. While in the Mandarin neutral tone case, the muscles that become lax are probably the laryngeal ones; in the case of English schwa reduction, they are the supra-laryngeal ones.

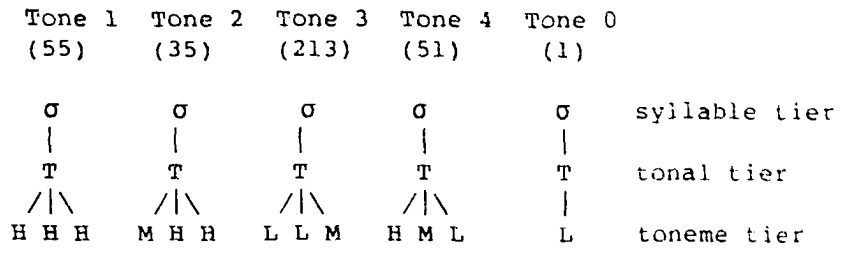
Although the precise phonetic value of the schwa may vary from environment to environment due apparently to co-articulation or mutual spreading between adjacent segments, it is still describable in terms of a single segment, a mid and lax vowel called schwa. A similar situation, I believe, is true in the case of the Mandarin neutral tone: no matter what the precise measurements of its pitch values, the fact that it is a short and lax low tone remains. It is only when this true nature of the neutral tone is securely pinned down can its various characteristics be understood.

Notice how the neutral tone case is elegantly accounted for by the present analysis along with all the other tone sandhi processes. In addition, it is analysed with virtually no further stipulation of rules. Notice also how the representation in the present analysis of Tone 0 in terms of a L toneme takes into account not only the pitch value of Tone 0, but also its unusually short duration, which has been unanimously observed.

6.3 *Summary and Conclusion*

This section contains first of all a brief summary of the analysis of Mandarin tone and tone sandhi proposed so far in this dissertation. First, the underlying representation of Mandarin lexical tones, and that of the neutral tone:

(6.48)



The characteristics of the representation are summed up as follows. In the underlying representation, (a) three levels of pitch contrast are assumed; (b) three tonemes are assumed to make up each of the four lexical tones; (c) the TBU is the syllable; (d) each tone is also represented as a unit in terms of a node T as well as a sequence of level tonemes; (e) each toneme accounts, in a relative sense, for 1/3 of the timing of the tone it belongs to by virtue of the following condition:

(6.49)

$$\tau(\text{toneme}) = 1/3\ \tau(\text{tone}) = 1/3\ \tau(\text{syllable})$$

where τ = time/duration

This condition serves (a) to indicate the durational relation between the T node/σ node and its tonemes, and (b) to limit the occurrence of the tonemes to no more than three in number for any given tone.

Under such a representational framework, all the tonal processes including Mandarin 2TS, 3TS(A), 3TS(B), 4TS, and 0TS can be uniformally addressed by the following Tone Reduction Principle:

(6.50). Tone Reduction Principle

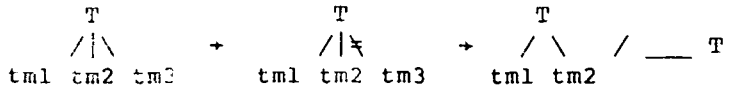
Clause A:
In normal speech, reduce a tone by one toneme iff it is immediately followed by another tone within the same prosodic foot.

Clause B:
In fast speech, reduce a tone by one toneme iff it is immediately preceded by another tone, and at the same time immediately followed by another tone within the same prosodic foot.

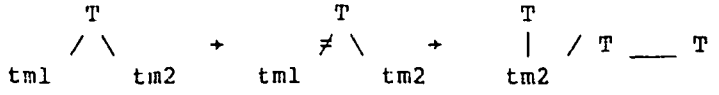
This Tone Reduction Principle functions to cut a tone in duration from its edges inward by the following implementation rules:

(6.51).

a. L-TTR



b. R-TTR

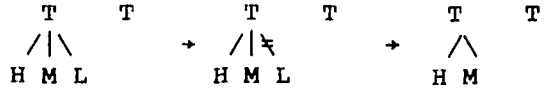


The Tone Reduction Principle is found to work in all of the tone sandhi processes[19] so far observed in Mandarin as well as the processes concerning Tone 0.

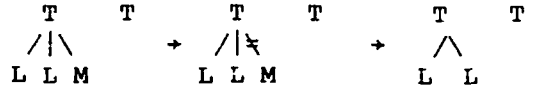
Specifically, the following derivations account for these processes respectively:

(6.52).

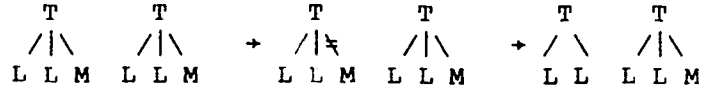
a. 4TS



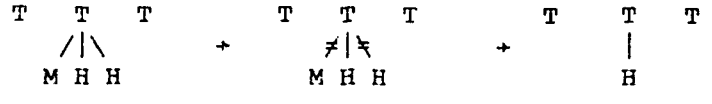
b. 3TS(A)



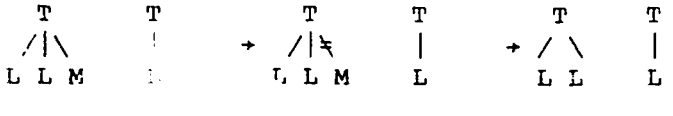
c. 3TS(B)



d. 2TS



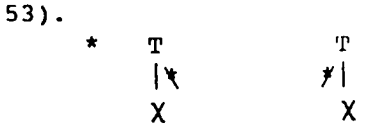
e. 0TS



Except for (d), which needs both Clause A and B of the Principle for its derivation, all the above processes are accountable in terms of the single L-TTR. Through the above illustrations of these processes, one can see that two of these processes distinguish themselves from the rest. Whereas the L-TTR produces grammatical results in all other cases, it merely yields OCP violations in the cases of 3TS(B) as in (c) and 0TS in (e).

Although the distinctiveness of the 3TS(B) process from other previously observed Mandarin tone sandhi processes has been much observed, no principled explanation has been provided for this distinctiveness. Under the present theory, it becomes explicable in a principled way. While all the other previously-observed tone sandhi processes are of the type triggered by the speech tempo, 3TS(B) has yet to satisfy OCP. The Obligatory Contour Principle consists in the following WFC for the case of the Mandarin tone:

(6.53).

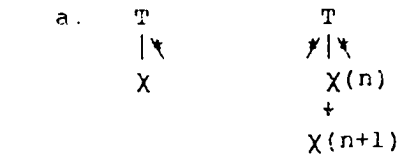


where X = H, M, or L

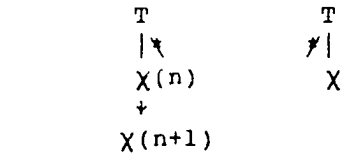
With this WFC in place, the two violations (shown in 6.49c and e) obligatorily undergo the following OCP repairs.

(6.54).

a.



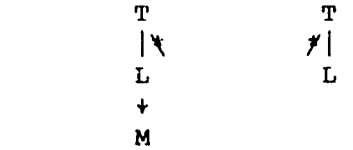
b.



where X = H, M, or L
and n = pitch level
thus for L, n = 1
M, n = 2
H, n = 3

The result of the 3TS(B) after L-TTR, meeting the structural description of the (b) rule in the above, is repaired by it through the following derivation,

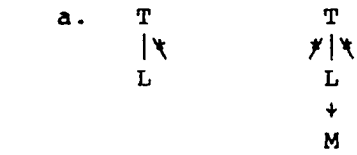
(6.55).



and that of 0TS, meeting the structural description of the (a) rule, is altered by it in the following manner:

(6.56).

a.



This way, the odd behavior of a low Tone 3 BEFORE A TONE 3 and that of a low Tone 0 AFTER A TONE 3 are elegantly related to each other.

6.3.1 Toward A Generalized Theory

To make a further summary of the framework just laid out, let me point out that the TTRs implementing the Principle may actually be collapsed together and stated as a more general rule. This rule may be stated in the following simplistic manner:

(6.57). Tone Reduction Rule (generalized)

Cut the adjacent toneme off the adjacent tone.

Of course, stated in such a general way, this Tone Reduction Rule must be constrained by some conditions. One such condition stands out as fairly obvious. Since utterance-final tones are not touched by any of the sandhi rules discussed, a filter seems needed to ensure just that. Here, the concept of extra-metricality laid out in non-linear phonology seems to come in handy:

(6.58). Condition of extra-metricality:

Domain-final tones in Mandarin are extra-metrical (and therefore insensitive to any sandhi rules).

With both (6.57) and (6.58), Y in a sequence of tones XY will cause X to loose its final toneme, but not vice versa. This would yield the results of 4TS, 3TS(A) and 3TS(B) at the stage before the OCP-repair processes. On the other hand, tone Z in a sequence of tones XYZ will cause its adjacent tone Y to loose its final toneme, tone Y will do the same to tone X, and tone X will cut off the initial toneme of the adjacent tone Y (which is not in final position). The last tone, tone Z, is however, not changed. The result is just as expected for 2TS, the first tone has two tonemes, the middle one, and the final three.

It should be pointed out here that lying at the basis of the present analysis of tone sandhi is the assumption that the surface shape of a tone is subject to timing. Specifically, the less the time afforded, the shorter the tone. Two factors are seen as controlling the relative timing:

(6.59). a. rate of speech
b. stress pattern

Stress pattern is a major determining factor for toneme reduction. The less the stress, the smaller the number of tonemes. A clear case of the stress-related temporal control is the appearance of the one-tonemed neutral tone: it may be one-tonemed even when it appears utterance-finally. Furthermore, the weaker the stress, the more likely the syllable that bears it will have a sandhi tone. Thus, in a two-syllable foot, the first tone is affected since the normal stress for a two syllable sequence is weak-strong. In a three syllable foot on the other hand, the medial syllable whose tone is bound to be affected is the weakest in stress among the three syllables.

Normally, however, it takes both rate of speech and stress pattern for the timing related type of tone sandhi (as against OCP-related dissimilatory tone sandhi) in Mandarin to occur. In these sandhi processes, the stress pattern identifies the site where the sandhi tone lies within the prosodic foot, while the rate of speech determines whether the sandhi process will occur. This means that the stress pattern is a necessary but insufficient condition for Mandarin tone sandhi to occur. It takes the rate of speech to actually activate the sandhi processes.

Though the precise measurement of the time allotted to a tone in proportion to that of another adjacent tone is probably hard to pin down (it would be irrelevant anyway at the phonological level), the timing in the relative phonological sense seems still capturable. In particular, when an sequence of, say, two syllables is uttered at a rather slow and deliberate speed, both tones involved are expected to be fully realized, and no sandhi should occur. As the speed in uttering this sequence begins to pick up, the first of the two tones is expected to be shortened with its last toneme deleted. Proceeding from there, one should soon expect further toneme loss on the part of the affected initial tone, when the rate of speech

accelerates further and the utterance becomes characterizable as fast speech. Given such expectations, it seems obvious that the Generalized Tone Reduction Rule should be allowed to occur cyclically. The cycles may perhaps be preliminarily stated as follows:

(5.60).

> In normal stress pattern, the operation of the Generalized Tone Reduction Rule at each cycle is determined by the rate of speech: the first cycle is effected at normal rate of speech; the second in fast speech.

In fact, there may exist a third cycle, whereby the last toneme left of the tone is reduced so that the tone becomes completely toneless. I will come to this point again later. Thus, eventually, as the cycle repeats, two situations may be the final results. One, the affected tone becomes contourless when only one toneme is left behind; two, the tone is reduced to zero or toneless when no toneme of the original tone exists any more. Clearly, the latter situation implies the former. In a crude way, I may perhaps capture the two situations in the following context-free rules:

(6.61). a. T → contourless
b. T → toneless

These rules will be further discussed in the next Chapter.

So far in this chapter, I have laid out an analysis of Mandarin tone sandhi, based on the tonal representational scheme proposed in Chapter 4. The core of the analysis is the Tone Reduction Principle. As has been demonstrated, the assumption of this principle yields for the first time a unified treatment of all of the major tonal processes well-observed in Mandarin. To conclude this chapter, I would like to point out a confusion in the analyses of Mandarin tones. Under the present analysis, one can clearly see that there are truly two quite distinct types

of tone sandhi processes on Mandarin tones, one due to a timing requirement, and the other to dissimilation requirement. Both, however have been treated with the same theoretical mechanism, segmental or non-linear. Such an undiscriminating way of treating these tonal processes cannot but overlook important insights into these processes.

In fact, it is probably correct to say that there are actually only two real tone-sandhi processes in Mandarin: 3TS(B) and 0TS. Of course, whether this claim is justified depends on how one defines the term "tone sandhi". If the term is defined as it commonly is, that the value of a tone alters because of the presence of the value of another tone in its periphery, then only these two cases of 3TS(B) and 0TS are true tone-sandhi cases. This is because only these two involve the alteration of the tonal value of a tone because of the tonal value of the tone before/after it. Since in all the other cases, it is the mere presence of a tone in the periphery and not the value of the affecting tone that brings about the alteration of the value of the affected tone, they probably should not be regarded as tone sandhi cases proper under such a definition of tone sandhi.

6.4 Notes to Chapter 6

1 To facilitate the reading of these rules, each of the four sandhi rules is also written in (ii)s with tone names.

2 In this dissertation, the terms "utterance-final position," "final position" and "domain-final position" are used interchangeably.

3 It should be noted, however, that these two rules (as well as the two rules in (6.3)) may give slightly different readings of the process involved. While the rule in (b) marks clearly a process of insertion, the rule in (a) may or may not involve insertion: there may be merely a tone-replacement process indicated. This difference is, however, not crucial here.

4 Woo does not actually use such abbreviations for her representation of this form of Tone 3. Her original representation is in feature matrices (p. 45, see also § 2.7):

[+low] [+low] [-low, -high]

5 As will be discussed in the next chapter, this principle has operated in other Chinese languages as well.

6 By "all Mandarin tone sandhi processes", I mean all the phonological tone-sandhi processes, which affect all the syllables in the relevant contexts. These, however, exclude the morphologically conditioned tone sandhi processes discussed earlier in § 4.2.4, which affect only a few frequently occurring morphemes.

7 It should be noted that Shih's use of H, L, M is exclusively for her acoustic study just cited. It is not a phonological framework for tonal analysis.

8 In the remaining sections of this chapter, comparison will often be made between the analysis I have proposed and the previous analyses. However, among these previous analyses, only three have been found comparable as far as Mandarin tones are concerned. Others are either followers of these three, or in an incomparable framework (e.g. Wang 1967, cf. § 2.4).

9 No comparison will be made between the present analysis and Woo's with regard to 4TS, since Woo does not discuss this tone sandhi process.

10 See § 2.8.2 for an explanation of OCP.

11 This WFC will be discussed again later in this chapter when it comes to the neutral tone case.

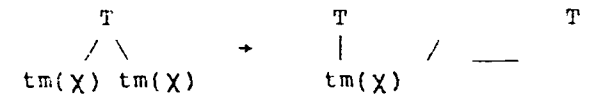
12 The merging process suggested here is further motivated by evidence from other Chinese languages to be discussed in Chapter 7, especially as it applies to level tones. However, intuitive and motivated as the process is, I have found it defying a reasonable formalization, at least within the non-linear theory the present analysis is couched in. What is needed is a formal phono-

logical rule which says the same thing as do the words in (6.26). However, it is not clear to me how this rule can be written within the present non-linear phonology (especially if it does not allow contour features). Due to the anticipated scope of this problem, I will not attempt to solve it here. One may in any event have to wait for further major advances in generative phonology in order to solve this problem.

13 I have, in fact, found an alternative analysis for 3TS(B) to the one I have just demonstrated, assuming still the Principle. Although, for the sake of uniformity of the analysis in accounting for all the sandhi cases addressed in this chapter, I have opted for the one just presented to be included in the package of the theory I have proposed, it is nevertheless harmless to present here the alternative approach. Future researchers might find some heuristic value in it.

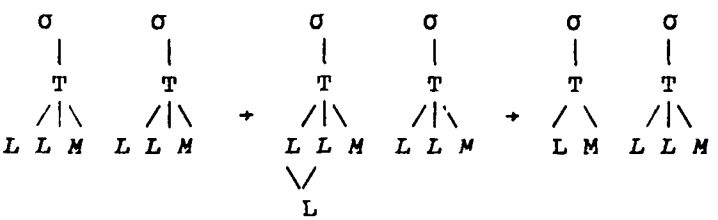
The existence of the alternative is a result of there being so many ways that the OCP can operate (Yip 1988). The analysis just proposed presumes one of these: the OCP works to fix rule output. In other words, it works after an OCP violation is produced by a rule. Although such an operation of the OCP is a legitimate one, the OCP has also been found to block the application of a rule which would otherwise produce a violation (Yip 1988). Suppose that the latter is in fact the case for 3TS(B), what then is the analysis? As seen in (6.20), TTR produces an OCP violation given (6.21). If the OCP has stepped in to block its application, what then is next? It seems that there must be an alternative rule to TTR, a rule that will perform the same function of tone reduction designated by the Principle, but at the same time, will not produce an OCP violation. If, as I have argued, the surface form of the rule is LM, this rule should eliminate a L toneme from the representation of LLM. I may perhaps formalize this rule as follows:

Tone Compression Rule (TCR)



What this rule says is that two continuous, identical tauto-tonic tonemes will be collapsed into one if their common mother node T immediately precedes another T. As I discussed earlier, the identical toneme clusters in Mandarin are true geminates. If so, a deletion treatment of one of the L tonemes is not desirable. In any event, it would be hopelessly arbitrary to say that either of the two L tonemes is the one deleted, and there is no good explanation for the deletion of a medial toneme in a tone in that context. On the other hand, it seems natural to see that a longer curved tone is compressed to a shorter curved tone when there is a lack of time.

Thus, given TCR, the process involved in 3TS(B) can be formalized as follows:



A comparable situation involving competing rules (as the TTR and TCR) is also discussed in a recent study by Chen (1991). In this study, Chen addresses certain tone sandhi processes in the New Chongming dialect of Wu. Chen finds in this dialect that there are three competing tone sandhi processes which should necessarily be interpreted as tone repair strategies, since they all can apply on the same sequence of tones, but yield different results. And just as TTR and TCR, these three strategies compete so as to yield results that conforming to certain Well-Formedness Conditions, among which is, not surprisingly, the OCP.

14 The stress pattern is shown here using the techniques of metrical phonology.

The basic idea is that the more the grids, the stronger the stress. See Liberman and Prince 1977, and Selkirk 1984 for details.

15 The flattening effect of the 2TS (or rather the flattening effect of the present analysis as a whole) will be further discussed in Chapter 7.

16 That is, Tone 3 in its non-final form.

17 Dow does not provide the unit of these numbers. Presumably, it is Hz of fundamental frequency.

18 It should be noted that I do not make the claim that this representation of Tone 0 appears exactly in the lexicon. The assumption I adopt is instead that this representation of Tone 0 appears after morphological or syntactic stress assignment.

19 See Note 6 to this chapter.

Chapter VII

FURTHER EVIDENCE

7.1 Tempo and Domain Size: Cheng's Study

Earlier, in Chapter 1, in an attempt to define the focus of this dissertation, I mentioned that the present study would not address specifically Mandarin tone sandhi on a larger-than-the-minimal domain of the foot defined by prosodic phonology. However, the results from the analysis proposed in the present study do seem to relate to certain findings made in studies within prosodic phonology. One such finding is that the domain of Mandarin tone sandhi is sensitive to speech tempo (Cheng 1973, Shih 1986). According to Cheng, "when the speed increases, the size of the domain [for tone sandhi] increases (p. 53)." Below is a classic case first discussed by Cheng (p.52) to indicate among other things the tempo-sensitivity of Mandarin tone sandhi.

(7.1). [[lao li] [mai [hao jiu]]. "Old Li buy good wine."

UR:		3	3	3	3	3
SR:	a.	(2	3)	3	(2	3)
SR:	b(i).	(2	2	3)	(2	3)
	(ii).	(2	3)	(2	2	3)
SR:	c.	(2	2	2	2	3)

The above example contains a complete sentence with five third-toned monosyllabic words. The numbers 3 and 2 are short for Tone 3 and Tone 2 respectively. Indicated in this example is the fact that in normal or slow speech, the 3TS(B) process (however defined) works only on a domain of a minimal size (a). As the rate of speech increases, however, the domain becomes increasingly larger. Eventually, it covers the maximal domain of the sentence (c).

Why, then, does speed matter at all in the determination of the domain size? This question has not been addressed previously. It is not clear to me, anyway, how this question could be answered in a principled way by previous analyses in which this tone-sandhi process is considered to involve merely the change of pitch contour and not duration. However, when this tone sandhi is regarded as involving primarily a shortening process, a process of tone truncation as understood in the analysis I have proposed, the effect of speech tempo on the size of the domain becomes explicable.

The explanation is that although in normal speech, tone sandhi only affects a tone within a minimal domain, as the speed increases, what was realized to its full length of three tonemes (by being outside of the minimal domain) gets less and less temporal space, and up to a certain point of the speech tempo, it becomes shortened to the extent that its last toneme is deleted; it thus becomes "pressed" into an existing smaller domain either to its right or to its left (depending on the syntactic and/or prosodic structure), and thus loses its status of being "domain-aloof." The net result is that the original minimal domain is extended for the sandhi process.

So far in this dissertation, I have argued for my analysis with synchronic evidence. However, this analysis is not limited to the accommodation of synchronic tonal data, but it seems to have diachronic implications as well. If synchronically, a single-syllable morpheme, no matter in what tone underlyingly, should appear with a one-toneme tone in a fixed morphological (or syntactic) environment, one should expect that in the long run, this morpheme may turn out to become a level-toned morpheme. This is exactly what has been found in Chen's (1989) and Shih's (1986) studies, which will be discussed in the remainder of this chapter.

7.2 Diachronic Evidence: Chen's Study

Chen's (1989) study contains a historical account of certain tonal changes in Mandarin. Specifically, her study examines tonal changes in the reduplicated forms in Mandarin. By comparing entries of reduplicated forms in four dictionaries representing two separate points of time, 1932 and 1963 to be more specific, Chen finds that in a reduplicated word of shape of XYY (the symbols X and Y each represent an individual syllable), the YY part may appear as any of the following five tonal shapes, YY, Y0, 00, 10, 11 (Y represents a syllable with its original tone; 0 represents a syllable with a neutral tone; and 1 represents a syllable with Tone 1). Similar changes are found, though less frequently, in the YY part of the reduplicated words of the shape XXYY. The following are some examples from her data:

(7.2).

	tones of YY		
X Y Y	origin	1932	1963
a.xuě lin lin	2 2	0 0	1 1/2 2
b.bái mang mang	2 2	2 2	1 1/2 2
c.lǎn yang yang	2 2	2 0	1 1
d.chì luo luo	3 3	3 3	1 1/3 3
e.zhí ting ting	3 3	3 3	1 1/3 3
f.míng huang huang	3 3	0 0	1 1/3 3
g.hēi dong dong	4 4	4 4	1 1
h.chén dian dian	4 4	4 4	1 1
i.xiāng fu fu	4 4	4 4	1 1

The numerals 1, 2, 3, 4 and 0 in the above stand for Tone 1, 2, 3, 4, and 0 respectively. The slashes are used to mean either the numbers on the left or the numbers on the right, or both. That is, they are used in an inclusive sense. Also keep in mind that though written apart in the table, the three syllables XYY form one single word. The separate arrangements of them in the table is simply intended to facilitate an analytic reading. What is illustrated in the above shows that the reduplicated syllables (i.e. the YYs) of these three-syllable words have all undergone a change in their pitch values ultimately to that of Tone 1, regardless

of the lexical tones these syllables may originally have had. One characteristic of these three-syllable reduplicated words is that they can all be characterized in term of a dactylic rhythm, as shown below:

(7.3).

X	Y	Y
long	short	short

In fact, in terms of duration, the Ys can be said to each contain one toneme.

Now let us consider what has caused the diachronic change. My explanation is that these Ys have been one-tonemed in those fixed morphological positions for a long-enough period of time. Namely, the Y morphemes, having appeared one-tonemed, and therefore, contourless, in their stable morphological positions for a historically long-enough period of time, began at some point to lose totally their original tonal identities, and became to be regarded as level tones for their own sake. Eventually they merged with the only level tone, Tone 1, in this language.

According to Chen, these Y syllables have gone through the following stages of change:

(7.4). Original lexical tone → Tone 0 → Tone 1

This proposed path of tonal change coincides precisely with the prediction of the change made by the analysis I have proposed. The original tone, having become one-tonemed, contourless neutral tone due to a shortage of timing, eventually merges with the only level tone, Tone 1, in the language. In a sense, if Chen's path serves to identify the direction and the stages of the change, my analysis confirms it by further explaining why the direction and the stages of the change. This is captured below in (a) alongside of Chen's path in (b):

(7.5).

a. Original lexical tone → one-tonemed
→ level tone (merging into Tone 1)

b. Original lexical tone → Tone 0 → Tone 1

Notice that there is a minor difference between my theory and Chen's: my theory entails that the path of change may, but does not have to, include the intermediate stage of Tone 0. The reduplicated Y syllables, occurring in the durationally-short positions in the dactyl, may change right into a Tone 1 given enough time in history, since the one-tonemed stage is also in itself a level-tone stage which can theoretically merge into Tone 1 all on its own. This may perhaps explain the observation made by Chen that for many of these three-syllable units, no documentation can be found to show that they were in the neutral tone at any point of time in the course of the evolution, or that they have indeed gone through the neutral-toned stage prior to becoming Tone 1. This lack of documentation may suggest that there is a possibility that, at least with some of these observed cases, the tonal change has bypassed the Tone 0 stage and headed directly into the Tone 1 stage.

In any case, no matter what the exact path of development, with or without the Tone-0 stage, my theory would correctly yield a level tone, in this case Tone 1, as the ultimate result of the change.

7.3 Diachronic Evidence: Shih's Study

The result of the analysis of tone sandhi developed in this study also supports a theory of tonal development proposed by Shih (1986). As has been described previously, the theory I have proposed works to shorten tones synchronically within certain contexts so that they may surface phonetically one-tonemed, and therefore, contourless. Diachronically, these shortened contourless tones may eventually fossilize and become categorially level tones in some fixed morphological positions, as was shown in our discussion of Chen in the above section. In fact, these fixed morphological positions may be of a more general nature: they may be simply the non-final positions of a multi-syllable word.

Specifically, given the theory developed in this dissertation, it would be expected that in the long run, fossilization of the affected tones in the lexical words[1] may take place, the original tonal contours being lost for good. Two logical outcomes may be expected from the fossilization:

(7.6). a. The affected tones may become fixed level tones.

b. More drastically, the affected tones may become completely toneless.

As will be shown below, both situations are found in the Chinese languages.

The losing for good of the original contours on the part of the affected tones may eventually bring about more drastic typological change of the tonal system of the language. In particular, it may eventually create new species of tones whose domain is no longer confined to that of a syllable but rather, covers the whole domain of the lexical word. Exactly this theory has been proposed previously in Shih (1986).

7.3.1 Two Types of Tonal Systems

In her 1986 Ph. D. dissertation, Shih develops a quite innovative theory of tone evolution. In the development of the theory, she first proposes an elegant topological taxonomy of the Chinese tones, based on a survey of tonal data from a variety of the Chinese dialects/languages. According to this taxonomy, the Chinese languages are classified into two types, syllable-tone languages and word-tone languages.[2] In a syllable-tone language, tone melodies are typically found within the domain of a syllable, while in a word-tone language, tone melodies are typically found over the domain of a lexical word. An example of the former is Mandarin, and examples of the latter are Shanghai and Mende (an African tone language of the Menda family).

In Mandarin, Shih further explains, there is an inventory of four tone melodies (i.e. four tones), and each Mandarin syllable is allowed one tone melody from the inventory. No tone melody exceeds the domain of a syllable. On the other hand, a word-tone language like Mende has a set of five tone melodies (or five tones), but the domain of a tone melody is the word, irrespective of the number of syllables in the word. Since the Mandarin tonal system is already a familiar case to us, I will only provide the relevant Mende data here. The data are cited on p.39 in Shih.[3]

(7.7).

	Tone Melodies	*1-Syllable words*	*2-Syllable words*	*3-syllable words*
(a)	H	kó	pélé	háwámá
(b)	L	kpà	bèlè	kpàkalì
(c)	HL	mbû	ngílà	félàmá
(d)	LH	mbǎ	fàndé	ndàvúlá
(e)	LHL	mbâ	nyàhâ	nìkílì

As shown above, the five Mende tones may each be realized on a word of one, two or three syllable(s). For example, the tone HL is found "squashed" on a word of one syllable *mbû*, "stretched" over a word of two syllables *ngílì* and spread out in some designated manner over the domain of a three-syllable word *félàmà*.

According to Shih, although there are extreme cases like Mende and Mandarin, most of the Chinese languages/dialects contain more or less both types of tone melodies, and thus, fall somewhere in between the two extremes. Together, she maintains, these languages/dialects form a continuum. Given below is a simplified version of this continuum (see Shih, p. 83, for details):

(7.8).

where σ = pure syllable-tone languages/dialects
μ = pure word-tone languages/dialects
MSD = Mandarin as Standard Dialect
SH = Shanghai

7.3.2 The Historical Development

From the macroscopic viewpoint offered by the continuum, Shih envisages a developmental tendency in these languages, a tendency which, she says, *neutralizes* syllable tones and gradually[4] turns a language from a pure syllable-tone language via an intermediate stage where both types of tones co-exist,[5] to eventually a word-tone language when the syllable tones give way completely to word tones.

Now, let me digress to discuss a historical change in the Chinese language much related to the present discussion, in order to further facilitate understanding of the shortening theory under discussion. A well-attested theory about the Chinese languages is that they were once one monosyllabic language with almost exclusively monosyllabic lexical items. However, history has since witnessed a gradual increase in the number of multi-syllabic lexical items in the language (Norman 1988). Hence, today, the vast majority of Chinese lexical items[6] are multi-syllabic (most of which are dissyllabic compounds).

This situation can be seen by examining entries in any contemporary Chinese dictionary, in which multi-syllabic entries far exceed monosyllabic ones in number. Examples of such dictionaries are Xiandai Hanyu Cidian [Contemporary Chinese Dictionary] (1979), Xinhua Zidian [New Chinese Dictionary] (1985), Han-Ying

Cidian [Chinese English Dictionary] (1982), and Xinbian Xuesheng Zidian [Newly-Compiled Student Dictionary] (1989), to name just a few.

Not only has the number of multi-syllabic words increased greatly in the Chinese languages, cohesiveness among the components of such words has also been getting stronger. For example, Norman (1988), talking about the development of compounds in the language, remarks:

> "The degree of cohesion between the elements of a compound in the early period was considerably smaller than it became in later times. A compound term like *shīlǚ* 'troops' consists of two independent free morphemes, *shī* 'a military contingent of 3,500 troops' and *lǚ* 'a military contingent of 500 men'... This is quite different from the modern equivalent of the term, *jūnduì*, where neither *jūn* 'military, army' nor *duì* 'ranks, group, crew' is syntactically free; both occur only as members of compounds, and cannot function alone as independent nouns or adjectives." (p.86)

What has this development to do with the present discussion of tonal change? Well, it is obvious that a word tone (or melody) found over the domain of a multi-syllable word is only a possibility when there do exist multi-syllabic words; on the other hand, no such tones should have been possible had the language remained monosyllabic. Thus, what multi-syllabification in the Chinese languages does is to yield a necessary condition for the development of word tones.

It should be pointed out that this theory that multi-syllabification of Chinese lexical items provides the necessary condition for the projected tonal change is not my invention, I have only further explained it. It is first suggested by Shih. Specifically, Shih maintains that, as the process of multi-syllabification proceeds,

"the functional load of tones is reduced..., and at this stage, tonal neutralization is made possible (p. 65)."

Now, let us return to Shih's theory of the tonal development by first capturing roughly the essence of it as follows:

(7.9). syllable-tone language → word-tone language

One of the major contentions of Shih is that all tone sandhi processes can be seen as being motivated by this historical transition:

> "The typological change from syllable-level melody to word-level melody is a general tendency shared by many languages, and I would argue that it is the underlying force that motivates the various tone sandhi rules described." (p. 84)

Though she does not provide a detailed analysis to substantiate this theory of hers, she nevertheless provides much evidence to show that the tone sandhi processes in Chinese serve to neutralize contrasts among syllable tones. Neutralization, according to her, means the reduction of the number of possible tonal combinations. For instance, if a language has three tones, it would have 9 (or 3X3) possible tonal combinations on dissyllabic words. However, due to the historical neutralization process, that number 9 may be reduced to, say, 6.

I find that analysis of tone sandhi developed in the present study entails exactly such a neutralization function. In Mandarin, there are a total of 64 (or 4X4X4) tonal combinations on three-syllable words. The tone sandhi rules of my analysis would reduce that number to 48 (or 4X3X4) if the affected medial tones are shortened to one-toneme (and if they are not merged together yet). Otherwise, the number may be reduced to as small as 16 (4X1X4) if the affected medial tones all merge together to become one and the same level tone.

Similarly, there are a total of 4X4 = 16 possible tonal combinations for two-syllable compounds in Mandarin if those ending in the neutral tone are disregarded. In very fast speech, that number might be reduced to 3X4 = 12 when the affected initial tones are neutralized to only three possibilities of register height - H, M, or L (and if the affected tones are only shortened and not merged together yet). Otherwise, the number may be as small as four if the affected initial tones are all merged together. The above discussion implies three possible outcomes from the historical neutralization process under discussion (cf. 6.58, and 7.6).

(7.10)

a. (i). The original tones lose their contours but not relative register height.

(ii).The original tones lose their contours and merge together into one type of level tone.

b. The affected tones become completely toneless.

As I will show later, all three kinds of results are found in the Chinese dialects/languages.

It seems that the tonal theory I have developed provides further substance to Shih's insightful but somewhat skeletal theory. As can be expected of any historical development of such a magnitude, the whole transition from a syllable-tone language to a word-tone language should be extremely complex. A number of processes may be discernable along the way; together these processes may be held responsible for the completion of the transition. Moreover, these processes may interact in all ways possible. They may, for example, either work simultaneously in harmony or competing with each other, or they may work in chronological order. Also, these processes may be subject to various constraints or filters of either a language-universal or language-specific nature.

Whatever the true face of the transition, one process that does seem to have a place in it is the neutralization process. However, it seems that even the neutralization process as a whole has internal complexity. As is implied in Shih (cf. Shih, Chapter 4), there may exist several types of neutralization processes. One neutralization process discussed by Shih, for instance, refers to the historical process which has reduced the last syllable of many Mandarin dissyllabic compounds to the bearing of the neutral tone. This neutralization process may be differentiated from the type of neutralization process predicted by the theory I have developed in that the former is primarily stress-induced while the latter relates mostly to speech tempo.

Thus, it seems that the neutralization process understood in my analysis of the Mandarin tone sandhi is yet another type of the presumably several kinds of neutralization processes, and forms an integral part in the whole transition from a system of syllable tones to one of word tones. Thus, to avoid confusion in the exposition that follows, I will hereafter refer to the neutralization process projected by my analysis as a *flattening process*. The term "flattening" implies that the tones going through the process all become contourless, (mergeable into the same level tone or tones) or toneless. In a rather general way, the process can be captured by the following rules:

(7.11). a. T → contourless
b. T → toneless

Notice that this flattening analysis is just the diachronic version of my synchronic analysis of the tone sandhi processes in Mandarin. In what follows, I will discuss relevant data from the Chinese languages other than Mandarin to demonstrate evidence in support of the flattening theory. These data are primarily those presented in Shih in the development of her theory of tonal evolution (cf. Shih, Chapter 4, pp. 66-84). In discussing Shih's data, I will also attempt to give a more detailed interpretation of them in light of the flattening theory I have proposed.

7.3.3 Evidence from Xianyou

Earlier in this chapter, I pointed out that the synchronic flattening process (i.e. the process effected by the Tone Reduction Principle), working on a dissyllabic word, may in the long run lead the affected tone of a dissyllabic word to become "permanently" contourless. Evidence for such a change has been found in a Min dialect called Xianyou (cited in Shih, p. 71, from an earlier study by Dai 1958).

Xianyou has been found to have much more extensive tone sandhi processes than Mandarin. One of the fundamental difference between Xianyou and Mandarin tone sandhi is that in Mandarin, a derived tone is normally readily traceable to its underlying tone, and there is clearly only one underlying tone for any derived tone, whereas in Xianyou, a derived tone can have many underlying citation tones.

According to Shih, Xianyou has seven citation tones, which together allow a total of 49 (7X7) dissyllabic combinations. However, as she further points out, only 23, or about half of the 49 combinations, do actually surface. The following Xianyou data are from her (p. 71):

(7.12).

\ Y / X \	55	24	33	54	41	21	32
55						24 Y	
33							
41		33 Y					
24						41 Y	
21					55 Y		
32		54 Y					
54		32 Y					32 Y

In the above, the column on the extreme left are the X tones (the first tones) in the dissyllabic sequence XY, while the horizontal line on top lists the Y tones

(the second tones). In between the two axial lines are the combinations of the X and Y tones. Among the first features one would notice reading the above chart is the invariability of the Y tones, no matter in what tonal combination, the second tone of a dissyllabic word never changes (cf. the Condition of Extra-metricality in 6.55). Another thing one cannot fail to notice is of course the small number of tonal combinations.

Shih considers this much smaller number of the surface combinations to be due to a large-scale historical neutralization process whereby the initial tones in various tonal combinations are collapsed to but a few surfaces shapes. I believe Shih's interpretation is essentially correct. However, Shih does not provide a detailed explanation for the neutralization process. For instance, there is one remarkable detail in the Xianyou data that has not been explained in Shih. This detail concerns the fact that among the first tones that have merged together in the various combinations, about one half surfaced as the same level tone (33). Namely, five of the seven citation tones are flattened into the same level tone (33) when appearing as the first of the two tones. This part of the Xianyou data is shown again as follows:

(7.13). X Y → X' Y

a. 55 55 → 33 55
55 24 → 33 24
55 33 → 33 33
55 54 → 33 54
55 41 → 33 41

b. 33 55 → 33 55
33 24 → 33 24
33 33 → 33 33
33 54 → 33 54
33 41 → 33 41

c. 41 55 → 33 55
41 24 → 33 24
41 33 → 33 33
41 54 → 33 54
41 41 → 33 41

d. 24 55 → 33 55
24 24 → 33 24
24 33 → 33 33
24 54 → 33 54

e. 21 55 → 33 55
21 24 → 33 24
21 33 → 33 33
21 54 → 33 54

What is seen above is a process which may be roughly captured in the following rule:

(7.14).

or: T → 33 /  T

where w = a lexical word

How can one account for the existence of this rule and its wide application across tonal categories? This question, I think, should not be too difficult to answer if my theory of the flattening process is assumed. In light of the flattening theory, one can see that the tones of the first syllables in these relevant dissyllabic words were at one time uttered as contourless one-tonemed short tones due to processes similar to the temporal-controlled Mandarin tone sandhi processes dis-

cussed earlier. These short tones eventually converge into the same level tone (33). Thus, within my theory, the origin of these Xianyou tone sandhi processes are explained.

It should be pointed out, however, that although my theory of the flattening process covers a substantial amount of the data which amounts to about half of the surface forms of the various tonal combinations in dissyllabic words in the Min dialect of Xianyou, there are nevertheless many others which it does not account for. Although I will not attempt to explain those other cases, it is nevertheless harmless to speculate on the explanation that the answer must lie in the lexical diffusion theory of Wang (1969). According to this theory, it is possible that different neutralization processes may be at work simultaneously and competing with one another.

7.3.4 Evidence from Zhangping

Quite similar evidence as seen in Xianyou is cited in Shih from Zhangping, another Min dialect (Shih's data are from Zhang 1982a,b), as shown below:

(7.15).

\ Y / X \	24	11	55	53	31	21
24						
11						
21		33 Y			55 Y	
55						
31						
53		21 Y				

In contrast to Xianyou, Zhangping has six (rather than seven) citation tones, allowing for 6X6=36 tonal combinations on dissyllabic words. However, tone sandhi collapses all of these to virtually three types of combination, 33 Y, 55 Y, and 21 Y. Shih again correctly accounts for these data by maintaining that they are the result of a large-scale historical neutralization process. However, again there

seems to be more to this general observation. What then is the nature of this neutralization process? Can anything specific be pinned down? The answer to the latter question is yes. In fact, the Zhangping case is even more remarkable than the Xianyou one for reasons that will soon be made clear.

In Xianyou, only about half of the data are accounted for by the flattening theory. Here in Zhangping, all the data seem to fall out from it. Now let us have a closer look: all the initial tones, no matter in what position, fall into three and only three possibilities. But what exactly are these possibilities? I argue that they are three register tones distributed evenly over the pitch range, HH (55), MM (33), and LL(21). Representing (21) as LL should not be surprising given that it is normally the case that a phonological low tone has at the onset a slight fall (cf. the normal treatment of the third Mandarin tone discussed in § 4.2.3).

If the above understanding of the three derived tones is correct, the tonal phenomenon found in Zhangping dissyllabic words provides interesting evidence for the proposed flattening theory: all the initial tones of the two-syllable words in this dialect have become contourless, although they have not all merged into one and the same tone but are rather distributed evenly over the pitch range as three level tones. In fact, the observation that all the six citation tones merge into three different level tones rather than a single one may perhaps represent an earlier stage of the development. In any event, it is a logical outcome of the flattening process as described in (7.10a-i).

7.3.5 Evidence from Fuzhou

Fuzhou is another Min dialect, spoken in the capital city of the Fujian province. Like Xianyou and Zhangping, Fuzhou also shows extensive flattening of the first tones in dissyllabic words. The Fuzhou data given below are from Wright (1983, p.67, who in turn cites Chen and Norman 1965).

(7.16).

X \ Y	44	52	4	22	12	13	242
44							
12	44 Y			52 Y			
13							
242							
52		22 Y					
22/13				35 Y			

In contrast to Zhangping, Fuzhou shows evidence of a two-way merging of the initial tones in the dissyllabic words. In particular, these initial tones have merged either into (44) or (22). On the surface, it seems that only 22 of the 42 combinations listed fall into the two way partition, but a closer look at the other two alternatives, (52) and (35), leads to different conclusions. Now observe the Y tones after (44) and Y tones after (52). The former are all higher register tones (44, 55, 4), while the latter are all lower register tones (22, 12, 13, 242), at least at the tonal onset. This kind of distribution of the two sets of Y tones is highly unlikely to be a mere coincidence. Rather, it suggests a rule-governed lower level phonetic alternation: the high tone (44) remains high before high tones, but plummets in pitch (i.e. becomes 52) before tones whose onsets are of lower register. In brief, (52) is derived from (44) when appearing before a low tone. If this line of reasoning stands, the proposed flattening theory covers not just 22 of the 42 combinations, but 22+16 = 38 (or 90%) of the 42.

Similar arguments can be applied to the value of (35). It seems that this value is after all derived. Now observe the last line of the above chart. The observation is that the two X tones in this last line, (22) and (13), both low, appear as (22) before the higher register Y tones of (44), (55) and (4), but as (35) before lower

onset Y tones (22, 12, 13, 242). Such a tonal distribution is again to be unlikely an arbitrary result. This unlikelihood becomes even more clear when one realizes how natural it is for the derived X tone (22) to be altered again by being before other tones of its register height.

Unlike (52) which is derived from (44) by way of a spreading/assimilation process, here a dissimilatory process triggered by OCP is involved, just as that in 3TS(B) discussed in the last chapter. Specifically, the dissimilation rule can be loosely formalized as follows:

(7.17). 22 → 35 / ___ {22, 12, 13, 242}

or: LL → MH / ___ L

(cf. 3TS(B)).

If what I have argued for is correct, the Fuzhou dissyllabic tonal combinations show yet more evidence of the operation of a total flattening process. In particular, all the initial tones were flattened phonetically to level tones at one time, and then merged in two directions to eventually become either (22) or (44) phonologically.

7.3.6 Evidence from Suzhou

Suzhou (cited in Shih from Ye 1979a,b) belongs to another of the Chinese languages called Wu, spoken in the city of Suzhou in Jiangsu Province. So far, I have only cited evidence that shows the neutralization of the non-final tones in a two-syllable words. The Suzhou dialect presents evidence for the neutralization of the medial tones of words of more than two syllables. Now, see the following data from Suzhou (Shih, p. 73):

(7.18).

a.	(44)	(52)	(412)	(412)	→	(44)	(44)	(44)	(31)
	pe	tsɶo	si	ka		pe	tsɶo	si	ka
	sad	sad	world	world		"Les Miserables"			
b.	(13)	(13)	(13)	(44)	→	(13)	(33)	(33)	(31)
	ɦiɑ̃	mae	ɦy	tɕin		ɦiɑ̃	mae	ɦy	tɕin
	lamb	wool	wrap	scarf		"wrap scarf"			
b.	(13)	(52)	(3)		→	(13)	(44)	(22)	
	le	pae	za?			le	pae	za?	
	blue	treasure	stone			"sapphire"			

According to Shih, all the medial tones have lost their original contours and have basically become but two level tones: (44) and (33):

> "[Suzhou] tones in medial positions are partially neutralized: the melody contrasts are lost but the register contrasts are maintained. That is, all contour tones are simplified to level tones of the corresponding height. High register tones, (44), (52), (412) and (5) are neutralized to (44); lower register tones, (13), (31) and (3), merge to (33)." (Shih, p. 72)

If Shih's observation is correct, Suzhou presents further data for the flattening analysis I have proposed. So far, all the four Chinese dialects I have shown appear to give strong evidence for our flattening theory. In particular, they provide evidence to show that tones in the medial positions (or initial positions in the case of two-syllable words) may in the long run be flattened to level tones.

Recall that the flattening theory has yet another predicted outcome (7.10b), that is, the affected tones may eventually be completely reduced so that the syllables that bear them become underlyingly toneless. Logically speaking, a syllable can be regarded as underlyingly toneless if one can prove that its surface tone comes somehow from elsewhere than its original citation tone. There may be several sources from which this toneless syllable can acquire a surface tone, from

tones on its periphery by way of spreading, phonetic interpolation, or by default assignment. All these situations have been found in the Chinese languages. In the following, such evidence is adduced from Danyang and Wuxi.

7.3.7 Evidence from Danyang

Danyang, a northern Wu dialect, has been unanimously considered a language with word tones (Zhang 1989, Bao 1990, Yip 1989a,b, besides Shih 1986). According to Shih (citing from Lü 1980), there are six citation tones in Danyang, (11), (33), (55), (24), (3) and (4), among which the two short tones (3) and (4) only occur in syllables ending in a final stop. Examples of these six tones, cited in Zhang (1989, p.64), are given below:

(7.19).

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

Besides the six citation tones, there are also six word melody patterns in this dialect (data from Shih):

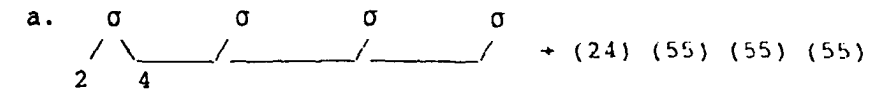
(7.20).

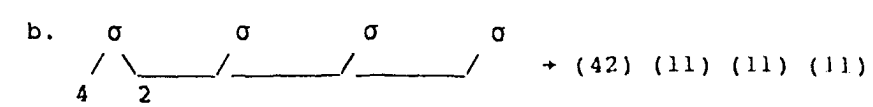
	dissyllabic	trisyllabic	quadrisyllabic
a.	(11) (11)	(11) (11) (11)	(11) (11) (11) (11)
b.	(33) (33)	(33) (33) (33)	(33) (11) (11) (11)
c.	(55) (55)	(55) (55) (55)	(55) (55) (55) (55)
d.	(24) (55)	(24) (55) (55)	(24) (55) (55) (55)
e.	(42) (11)	(42) (11) (11)	(42) (11) (11) (11)
f.	(42) (24)	(42) (42) (24)	(42) (42) (42) (24)

Two features are obvious concerning these word melodies. One, tones become identical from the second syllable on in cases of (a) to (e). Two, in the case of (f), tones found on the dissyllabic word are found at the edges of a longer domain (i.e. three or four syllable words), whose medial tones are identical to its initial tone. There have been two analyses for these melodies. One analysis (suggested in Shih)

assumes that the second element of the initial tone spreads to the rest of the tones. Let me take the trisyllabic words with the (d) and (e) melodies for an example to illustrate this analysis:

(7.21).

a.


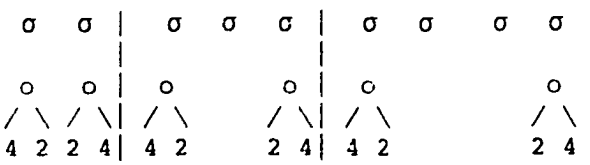
b.


Such an analysis covers all cases from (a) to (e), but not (f).

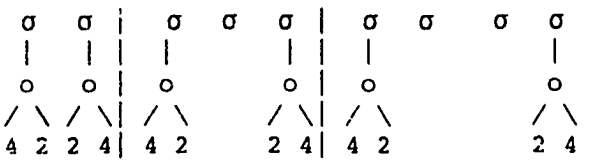
Another analysis is made specifically for cases such as (f). This is the analysis proposed by Yip (1989a,b). Motivating an edge-in type of association (rather than left-to-right or right-to-left association) for non-linear phonology, Yip argues that the (f) case is best analyzed in terms of association of the base melody shown on the dissyllabic word to the two edges of the domain, followed by the derivation of the medial tones through left-to-right spreading of the initial tone:

(7.22).

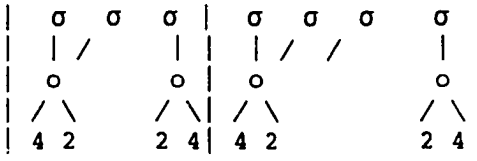
UR



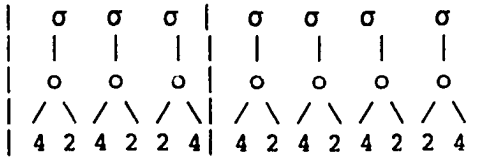
a. Association of Edges



b. Left-Right Spreading



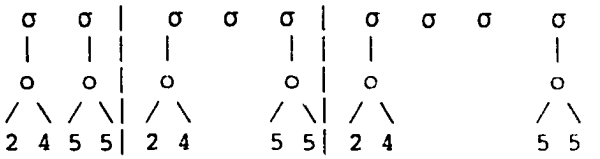
c. Tier Conflation



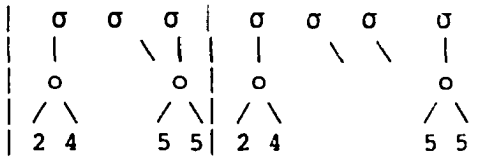
Obviously, Yip's analysis of the (f) case can be applied to the other five cases, if one wishes to achieve a unified analysis for all of the six word melodies. For instance, one can say that the base melody for (e) contains two tones: (42) (11). These two tones are anchored each to the two edges of the trisyllabic or quadrisyllabic word. This way, the difference between (f) and (a) to (e) becomes reducible to the difference in the direction of spreading (i.e. the second step in the whole derivation). While in the (f) case, the spreading is from left to right, in the rest of the cases, it is from right to left. Let me take the (d) case again to illustrate the latter kind of spreading:

(7.23).

a. Edge-in Association



b. Spreading



It seems that the application of Yip's analysis on the six cases (which is probably superior to the other analysis as it accounts for all the cases) differs from the other analysis in the treatment of the final tones. While with Yip's approach, the final tones are acquired through initial edge associations, in the other approach, they are derived through spreading. Despite the difference between the two approaches, one treatment that is consistent in both is that the medial tones are derived.

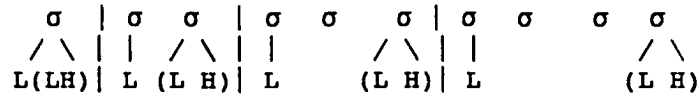
The fact that the medial tones are derived from tones on one edge of the domain presents clear evidence that the medial syllables are somehow toneless. Why are they toneless in those contexts? What happened to their citation tones that are still traceable in the dialect? These questions have not been answered. Within my theory, the answers are simple: the tones, appearing in the medial positions, are potentially to be flattened to toneless tones. This situation, is merely phonetic at first, but in the long run becomes more permanent and therefore phonological.

7.3.8 Evidence from Wuxi

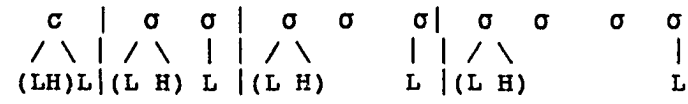
Wuxi is another Wu dialect that exhibits extensive use of word tones. According to Yip (1989a,b), Wuxi has four word-tone melodies: L(LH), (LH)L, L(HL) and H, each of which is found to flank both sides of a domain of a word (also see § 2.8.3). These melodies (Yip's data are from Chan and Ren 1986a,b) are described by Yip as follows: "the steep contours are fixed at the ends: A has a rise near the end, B has a rise near the start, and D has a fall near the end. Otherwise, the gradients decrease with the length of the domain (Yip 1989b, p. 157)." Pattern C, according to Yip, "is mainly high throughout (p. 159)," Having decided on the underlying melodic patterns (i.e. L(LH) for Pattern A, (LH)L for pattern B, H for Pattern C, and L(HL) for pattern D), Yip proceeds to an analysis of these patterns. Her analysis is given below:

(7.24).

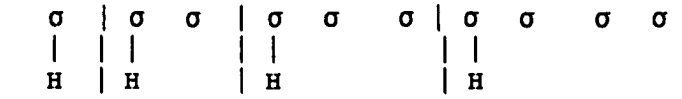
Pattern A:



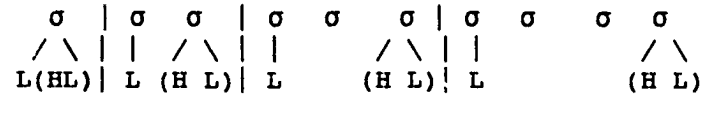
Pattern B:



Pattern C:



Pattern D:



Except Pattern C, Yip's analysis treats the basic tonal patterns seen in the monosyllabic words (i.e. the words in the left-most column) as decomposable into two parts, which are first of all fixed respectively to the two ends of a longer domain. This means that each of these three tonal pattern stretches further out to

fit an increasingly larger domain. No matter how big the domain, the two parts of the base are always split and anchored on its two edges respectively.

What is of the most interest here is, however, the medial tones with regards to where they are derived, for it is clear that tonal values of these medial syllables have nothing to do with their citation tones (no matter what their citation tones are). The actual tonal shapes of these medial syllables are in fact only definable with reference to the two tones on the edges. According to Yip, they are acquired through phonetic interpolation, or they together form a gradual transition of the pitch from one end to the other.

If Yip's theory is valid, there is yet further evidence to show that the affected tones, in this case affected medial tones, no longer carry any specific underlying tones in words, and the explanation for the lack of underlying tones is again attributable historically to the flattening process which at one time reduced these medial tones to toneless, a situation that in the long run is fossilized. These medial syllables, having lost their citation tones completely in their contexts, begin to acquire their surface tone values elsewhere. In the present case, these medial tones receive their surface tonal shape through phonetic interpolation between the edges of a domain of a word, the consequences of which are word tones.

7.3.9 Conclusion

In this section, I have ventured further into the realm of the diachronic implications of the theory of Mandarin tone sandhi I have proposed in this dissertation. The result of this venture gives a promising prospect (if not already a reality) of placing Mandarin tone sandhi in a more global context in the tonology of the Chinese languages as a whole. Guided by Shih's illuminating typology of syllable-tone versus word-tone system and her theory of the direction of evolution from the former to the latter, the venture has led to the increased visibility of a fine line in this developmental path broadly sketched by Shih.

What is seen in this fine line is a familiar case of a phonetically-conditioned synchronic change at the beginning of the path terminating in relatively permanent results where the previous phonetic conditions become no longer necessary. Mandarin tone sandhi, understood as being motivated phonetically in terms of timing of speech, gives substantial indication that Mandarin is still at the beginning of this developmental stage, a result consistent with Shih's theory (cf. 7.8). That Mandarin is still at the beginning of the developmental stage is indicated clearly by the fact that the change characterized by the flattening process (i.e. the Principle, from synchronic perspective) is still phonetically conditioned in this language; that is, all the individual tone sandhi changes effected by the rule are still synchronic processes.

Chinese dialects further down the developmental path, on the other hand, exhibit evidence indicating that the flattening process that is still operating in Mandarin phonetically has already produced fossilized results no longer conditioned phonetically. Evidence of such fossilization on initial tones of dissyllabic words are found in the abundance in the Min dialects of Xianyou and Zhangping, while evidence of such a diachronic change on the medial tones of words of more than two syllables is obvious in the Wu dialects of Suzhou, Danyang, and Wuxi.

And guided by Shih's diachronic theory, one may perhaps maintain further that the historical flattening process which starts synchronically on tones within certain prosodically defined contexts, as is the case in Mandarin, tends to reach dissyllabic lexical units first before extending to larger domains of lexical items. Thus, the Min dialects being on the middle reaches (cf. 7.8) of the development stream are marked by extensive flattening on the initial tone within the domain of a dissyllabic word. At this middle stage, the dialects are not yet confidently classified as word-tone dialects in view of an absence of more drastic neutralization involving words of greater length.

The Wu dialects almost at the terminating extreme of the development path show more complete tone flattening within their lexical items, no matter how long the lexical items are. At this stage, possible tonal combinations on multi-syllabic words are mostly flattened to but a few tonal melodies, with the underlying medial tones completely deleted and new tones from other sources distributed in their place. The net result is the emergence of true word tones.

7.4 Notes to Chapter 7

1 Lexical words are necessarily within tone sandhi domains (Shih 1986), especially for Chinese languages other than Mandarin.

2 Earlier in Chapter 4, I made a working distinction between the term "tone" for syllable tones and "melody" for word tones. As I explained there, this distinction is merely adopted to avoid unnecessary confusion in the exposition in this particular dissertation, and not treated as an absolute theoretical guideline. In this section, following Shih, the two terms may be used interchangeably when no confusion arises.

3 The Mende data were first described in Leben (1973). After being cited by Goldsmith (1979) as a strong evidence for autosegmental phonology, these data have appeared in numerous studies afterwards.

Following convention in the tonal studies of African languages, the diacritics ˊ, ˋ, ˆ, and ˇ are used exclusively here for these Mende data to stand for high, low, falling, and rising tones respectively.

4 Shih believes the change is in the manner of lexical diffusion (Wang 1969, Chen and Wang 1975 and many others). The theory of lexical diffusion was developed by Wang (1969), and it says basically that

> "phonological change may be implemented in a manner that is phonetically abrupt but lexically gradual. As the

> change diffuses across the lexicon, it may not reach all the morphemes to which it is applicable." (Wang 1969, p. 9)

For a detailed description of the theory, please refer to Wang himself.

5 According to Wang (Wang 1969), this stage should be marked by a competition for survival by both types of tones.

6 The exact number of multi-syllabic lexical items may vary from one Chinese language to another. For instance, it is said that Mandarin has more dissyllabic items than Cantonese. The difference is, however, insignificant and negligible in the present discussion.

BIBLIOGRAPHY

1. Anderson, Stephen R. (1978). Tone features. In Victoria A. Fromkin, (ed.), *Tone: A Linguistic Survey*. New York: Academic Press. pp.133-175.

2. Abramson, Arthur S. (1962). *The vowels and tones of standard Thai: Acoustical measurements and experiments*. Bloomington: Indiana Research Centre in Anthropology, Folklore and Linguistics.

3. Bao, Zhi-ming (1990). *On the nature of tone sandhi*. Doctoral Dissertation, MIT.

4. Baxter, William H. (1992). *A handbook of old Chinese phonology*. New York: Mouton.

5. Brown, J. Marvin (1965). *From ancient Thai to modern dialects*. Bangkok: Social Science Association Press of Thailand.

6. Chan, Marjorie K. M. (1991). Contour-tone spreading and tone sandhi in Danyang Chinese. *Phonology*, 8:237-259.

7. Chan, Marjorie K. M. and Hong-ming Ren (1986a). Wuxi tone sandhi: From last to first syllable dominance. Paper presented at *the 19th International Conference on Sino-Tibetan Languages and Linguistics*, Ohio State University.

8. Chan, Marjorie K. M. and Hong-ming Ren (1986b). Wuxi tone sandhi: From last to first syllable dominance. *UCLA Working Papers in Phonetics*, 63:48-70.

9. Chao, Yuen Ren (1930). A system of tone letters. *Le Maître Phonétique*, 30:24-27.

10. Chao, Yuen Ren (1968). *A grammar of spoken Chinese*. Berkeley, California: University of California Press.

11. Chao, Yuen Ren and Lien Sheng Yang (1947). *Concise dictionary of spoken Chinese*. Cambridge, Massachusetts: Harvard University Press.

12. Chen, Chung-yu (1984). Neutral tone in Mandarin: Phonotactic description and the issue of the norm. *Journal of Chinese Linguistics*, 12:299-333.

13. Chen, Chung-yu (1989). Lexical diffusion of a tonal change in reduplicates and its implications. *Journal of Chinese Linguistics*, 17:96-127.

14. Chen, Leo and Jerry Norman (1965). *An introduction to the Foochow dialect*. San Francisco: San Francisco State College.

15. Chen, Matthew Y. (1979). Metrical structure: Evidence from Chinese poetry. *Linguistic Inquiry*, 10:371-420.

16. Chen, Matthew Y. (1990). What must phonology know about syntax? In Sharon Inkelas and Draga Zec, (eds.), *the Phonology-Syntax Connection*, pp.19-46. Chicago: the University of Chicago Press.

17. Chen, Matthew Y. (1991). Competing repair strategy in tone sandhi. Paper presented at *the Third North American Conference on Chinese Linguistics*, May 3 - 5, 1991. Cornell University.

18. Chen, Matthew Y. and William S-Y. Wang (1975). Sound change: Actuation and implementation. *Language*, 51:255-281.

19. Cheng, Chin-chuan (1973). *A synchronic phonology of Mandarin Chinese.* The Hague: Mouton.

20. Cheng, Lisa L. (1987). On the prosodic hierarchy and tone sandhi in Mandarin. *Toronto Working Papers in Linguistics*, 7:24-52.

21. Chomsky, Noam and Morris Halle (1968). *The sound pattern of English.* New York: Harper and Row.

22. Chung, Raung-fu (1992). The domain of Hakka tone sandhi. Paper presented at *the Fourth North American Conference on Chinese Linguistics*, May 8 - 10, 1992. University of Michigan.

23. Clements, George N. (1985). The geometry of phonological features. *Phonology Yearbook*, 2:223-250.

24. Clements, George N. (1989). *A unified set of features for consonants and vowels.* Manuscript, Cornell University.

25. Clements, George N. and Samuel Keyser (1983). *CV phonology: A generative theory of the syllable.* Cambridge, Massachusettes: the MIT Press.

26. Dai, Qing-xia (1958). Tone sandhi patterns in the Wu dialect of Xianyou. *Zhōngguó Yǔwén*, 10:485-487.

27. DeFrancis, John (1984). *The Chinese language: Facts and fantasy.* Honolulu: University of Hawaii Press.

28. Dow, Francis D. M. (1972). *An outline of Mandarin phonetics.* Canberra: Australia National University Press.

29. Dreher, John and Pao-ch'en Lee (1966). *Instrumental investigation of single and paired Mandarin tonemes.* Douglas Advanced Research Laboratory.
:Duanmu, San (1990). *A formal Study of Syllable, Tone, Stress and Domain in Chinese Languages.* Doctoral Dissertation, MIT.

30. Elimelech, B. (1974). On the reality of underlying contour tones. *UCLA Working Papers in Phonetics*, 27:74-83.

31. Gao, Yu-zhen (1980). Běijīng huà de qīngshēng wèntí, [The neutral tone in the Beijing dialect]. *Yǔyán Jiàoxué yǔ Yánjīu*, 2:82-98.

32. Goad, Heather (1991). [Art] and [Rtr] are different features. Paper presented at *the Tenth West Coast Conference on Formal Linguistics*, 1991.

33. Goldsmith, John (1979). *Autosegmental phonology.* New York: Garland Publishing.

34. Goldsmith, John (1990). *Autosegmental and metrical phonology.* Cambridge, Massachusetts: Basil Blackwell.

35. Green, M. M. and G. E. Igwe (1963). *A descriptive grammar of Igbo.* Oxford: Oxford University Press.

36. Gruber, J. (1964). *The distinctive features of tone.* Unpublished manuscript.

37. Halle, M. and K. Stevens (1971). A note on laryngeal features, *Quarterly Progress Report*, 101, pp. 198-213. MIT.

38. *Hàn-Yīng Cídiǎn*, (1982). [Chinese-English Dictionary]. Beijing: the Commercial Press.

39. Harris, James W. (1983). *Syllable structure and stress in Spanish.* Cambridge, Massachusettes: the MIT Press.

40. Hayes, Bruce (1986). Inalterability in CV phonology. *Language*, 62:321-51.

41. Hockett, Charles (1947). Peiping phonology, *Journal of American Oriental Society*, 67:253-267.

42. Hockett, Charles (1955). A manual of phonology. *Journal of American Linguistics*, 21(4), Part I, Memoir II.

43. Howie, John Marshall (1974). On the domain of tone in Mandarin. *Phonetica*, 30:129-148.

44. Howie, John Marshall. (1976). *Acoustical studies of Mandarin vowels and tones.* Cambridge: Cambridge University Press.

45. Hung, Tony T. N. (1987). *Syntactic and semantic aspects of Chinese tone sandhi.* Doctoral Dissertation, University of California at San Diego.

46. Hyman, Larry M. (1985). *A theory of phonological weight.* Dordrecht: Foris Publications.

47. Itô, Junko (1988). *Syllable theory in prosodic phonology.* New York: Garland Publishing, Inc.

48. Jiang, Jian-ming and Yu-ling He (1987). *Yùnwén Gàilùn*, [An Introduction to Versification]. Shanghai, China: Gāodĕng Jiàoyù Chūbănshè, [High Education Press].

49. Jin, S. (1986). *Shanghai morphotonemics.* Bloomington: Indiana University Linguistics Club.

50. Kahn, Daniel (1976). *Syllable-based generalizations in English phonology.* Doctoral Dissertation, MIT.

51. Kaisse, Ellen M. (1985). *Connected speech: The interaction of syntax and phonology.* Toronto:Academic Press.

52. Kaisse, Ellen M. and Arnold M. Zwicky (1987). Introduction: Syntactic influences on phonological rules. *Phonology Yearbook*, 4:3-12.

53. Katamba, Francis (1989). *An introduction to phonology.* New York: Longman.

54. Kiparsky, Paul (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry*, 10:421-41.

55. Kratochvil, Paul (1968). *The Chinese language today: Features of an emerging standard.* London: Hutchingson University Library.

56. Krauss, Michael (ed.) (1985). Yupik Eskimo Prosodic systems: Descriptive and comparative studies. *Alaska Native Language Centre Research Paper, No. 7.*

57. Leben, William R. (1973). *Suprasegmental phonology.* Doctoral Dissertation, MIT. New York: Garland.

58. Leben, William R. (1978). The representation of tone. In Victoria A. Fromkin, (ed.), *Tone: A Linguistic Survey.* New York: Academic Press. pp.177-219.

59. Levin, Juliette (1985). *A metrical theory of syllabicity.* Doctoral Dissertation, MIT.

60. Li, Charles N. and Sandra A. Thompson (1981). *Mandarin Chinese: A functional reference grammar.* Berkeley: University of California Press.

61. Lieber, Rochelle (1987). *An integrated theory of autosegmental processes.* Albany, New York: State University of New York Press.

62. Lieberman, M. and A. Prince (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8:249-336.

63. Lin, Hua (1990). Toward a theory of Mandarin reduplication. *Working Papers of the Linguistics Circle*, University of Victoria, Canada, 9:129-141.

64. Lin, Hua (1991). The domain of third tone sandhi in Mandarin. Paper presented at *the 7th Northwest Linguistic Conference*, February 23-24, 1991, University of Victoria, Canada.

65. Lowenstamm, Jean and Jonathan D. Kaye (1985). Compensatory lengthening in Tiberian Hebrew. In Leo Wetzels and Engin Sezer (eds), *Studies in Compensatory Lengthening.* Dordrecht: Foris Publications.

66. Lǚ, Shu-xiang (1980). Dānyáng fāngyán de shēngdiào. xìtǒng, [The tonal system of the Danyang dialect]. *Fāngyán*, 1980(2):85-122.

67. Maddieson, I. (1972). Tone system typology and distinctive features. *Proceedings of the 7th International Congress of Phonetic Science*, pp.958-961. The Hague: Mouton.

68. McCarthy, John J. (1979). *Formal problems in Semitic phonology and morphology*. Doctoral Dissertation, MIT.

69. McCarthy, John J. (1982). Prosodic templates, morphemic templates, and morphemic tiers. In Harry van del Hulst and Norval Smith, (eds.), *the Structure of Phonological Representations*, Part I, pp.191-223.

70. McCarthy, John J. (1986). OCP effects: Gemination and antigenimation. *Linguistic Inquiry*, 17:207-264.

71. McCarthy, John J. (1988). Feature Geometry and Dependency: A Review. *Phonetica*, 43:84-108.

72. McCarthy, John J. and A. Prince (1986). *Prosodic morphology*. Manuscript.

73. McCarthy, John J. and A. Prince (1990). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory*, 8:209-283.

74. Norman, J. (1988). *Chinese*. New York: Cambridge University Press.

75. Nespor, Marina and Irene Vogel (1983). Prosodic structure above the word. In A. Cutler and D. R. Ladd, (eds.), *Prosody: Models and Measurements*, pp.123-146. Berlin: Springer-Verlag.

76. Nespor, Marina and Irene Vogel (1986). *Prosodic phonology*. Dordrecht: Foris Publications.

77. Odden, D. (1988). Anti antigemination and the OCP. *Linguistic Inquiry*, 19:451-476.

78. Ohala, John J. (1978). Production of tone. In Victoria A. Fromkin (ed.), *Tone: A Linguistic Survey*. New York: Academic Press. pp.1-39.

79. Packard, Jerome L. (1989). Register in Chinese tonal phonology. In Marjorie Chan and Thomas Ernst, (eds.), *Proceedings of the Third Ohio State University Conference on Chinese Linguistics*, 13-14 May, 1988, pp.18-36. Reproduced by Indiana University Linguistics Club.

80. Pike, Kenneth L. (1948). *Tone languages*. Ann Arbor: University of Michigan Press.

81. Pike, Kenneth L. (1967). *Language in relation to a unified theory of human behavior*. The Hague: Mouton.

82. Pulleyblank, D. (1986). *Tone in lexical phonology*. Dordrecht: Reidel.

83. Ramsey, Samuel Robert (1987). *The languages of China.* Princeton, New Jersey: Princeton University Press.

84. Qi, Sheng-qiao (1956). Hànyǔ de zìdiào, tíngdùn yǔ yǔdiào de jiāohù guānxi, [the interaction among Chinese tones, pauses and intonations]. *Zhōngguó Yǔwén*, 10:10-12.

85. Sampson, G. (1969). A note on Wang's phonological features of tone. *International Journal of American Linguistics*, 35:62-66.

86. Schuh, Russell, G. (1978). Tone Rules. In Victoria A. Fromkin, (ed.), *Tone: A Linguistic Survey.* New York: Academic Press. pp.221-256.

87. Selkirk, Elizabeth O. (1978). *On prosodic structure and its relation to syntactic structure.* Bloomington: Indiana University Linguistics Club.

88. Selkirk, Elizabeth O. (1984). *Phonology and syntax: the relation between sound and structure.* Cambridge, Massachusettes: the MIT Press.

89. Selkirk, Elizabeth O. (1986). On derived domains in sentence phonology. *Phonolcgy Yearbook*, 3:371-405.

90. Selkirk, Elizabeth O. and Tong Shen (1988). Tone deletion in Shanghai Chinese. In *Sentence Phonology*, Vol. 1. Amherst: University of Massachusettes.

91. Selkirk, Elizabeth O. and Tong Shen (1990). Prosodic domains in Shanghai Chinese. In Sharon Inkelas and Draga Zec, (eds), *the Phonology-Syntax Connection*, pp.313-337. Chicago: the University of Chicago Press.

92. Shen, Susan Xiao-nan (1990a). *Prosody of Mandarin Chinese.* Berkeley: University of California Press.

93. Shen, Susan Xiao-nan (1990b). Tonal Co-articulation in Mandarin, *Journal of Phonetics*, 18:281-295.

94. Shen Susan Xiao-nan (1991). A study of rhythm in Mandarin prosody. Paper presented at *the Third North American Conference on Chinese Linguistics*, May 3 - 5, 1991. Cornell University.

95. Shen, Tong (1985). The underlying representation of Shanghai tones. *Yǔyán Yánjiū*, 2:85-101.

96. Shen, Tong (1986). *The formation of tone groups in Shanghai.* Manuscript, University of Massachusettes.

97. Shen, Y. (1988). *A Tentative Hypothesis regarding trisyllabic tone sandhi in Pingyao.* Manuscript, University of California at San Diego.

98. Shih, Chi-lin (1986). *The prosodic domain of tone sandhi in Chinese.* Doctoral Dissertation, University of California at San Diego.

99. Shih, Chi-lin (1987). *The phonotactics of the Chinese tonal system.* Bell Laboratory technical memorandum.

100. Steinbergs, Aleksandra (1987). The classification of languages. In William O'Grady and Michael Dobrovolsky, (eds.), *Contemporary Linguistic Analysis: An Introduction.* Toronto: Longman Company.

101. Vogel, Irene (1984). On constraining prosodic rules. In Harry van der Hulst and N. Smith, (eds.), *Advances in Non-Linear Phonology*, pp.217-233. Dordrecht: Foris Publications.

102. Wang, William S-Y. (1967). Phonological features of tone. *International Journal of American Linguistics*, 33:93-105.

103. Wang, William S-Y. (1969). Competing changes as a cause of residue. *Language*, 45:9-25.

104. Wang, William S-Y. (1987). A note on tone development. *Wang Li Memorial Volume*, pp. 435-443. Hong Kong: Joint Publishing Co.

105. Wang, William S-Y. (1991). *Explorations in Language.* Taipei, Taiwan: the Pyramid Press.

106. Wang, William S-Y. (1991). (Ed.). Languages and Dialects of China. *Journal of Chinese Linguistics* Monograph Series No.3.

107. Wang, William S-Y. and Chin-chuan Cheng (1987). Middle Chinese tones in modern dialects. In R. Shannon and L. Shockey, (eds.), *To Honor Ilse Lehiste*, pp. 513-523. Dordrecht: Foris Publications. y

108. Woo, Nancy (1972). *Prosody and phonology.* Doctoral Dissertation, MIT. Distributed by the Indiana University Linguistics Club.

109. Wright, Martha Susan (1983). *A metrical approach to tone sandhi in Chinese dialects.* Doctoral Dissertation. University of Massachusettes.

110. *Xiàndài Hàiyǔ, I.* (1978). Compiled by Yāntái Shīzhuān, Tàiān Shīzhuān, Hézé Shīzhuān, Běizhèn Shīzhuān [Yantai Teacher's College, Taian Teacher's College, Heze Teacher's College, Beizhen Teacher's College], Shandong, China.

111. *Xiàndài Hàiyǔ Cídiǎn*, (1979). [Contemporary Chinese Dictionary]. Beijing: the Commercial Press.

112. *Xīnbiān Xuéshēng Zidiǎn*, (1989). [Newly-Compiled Chinese Dictionary for Students]. Beijing: Jindun Press.

113. *Xīnhuá Zidiǎn*, (1985). [New Chinese Dictionary]. Beijing: the Commercial Press.

114. Xu, Shu (1983). Qīngshēng de zuòyòng, [The functions of the neutral tone]. In Zheng-kun Wang, Wen-qing Xie, and Zhen-duo Liu, (eds.), *Yǔyánxué Zīliào Xuǎnbiān*, [Selected materials on linguistics], pp. 220-224.

115. Yan, Jing-zhu and Mao-can Lin (1988). Běijīng huà sānzìzǔ zhòngyīn de shēngxué biǎoxiàn, [The acoustic characteristics of the stress on the trisyllables of the Beijing dialect]. *Fāngyán*, 1988(3):227-37.

116. Ye, Xian-ling (1979a). Tone sandhi in Suzhou. *Fāngyán*, 1979(1):30-46.

117. Ye, Xian-ling (1979b). More on the tone sandhi of 'shang' and 'qu' tones in Suzhou. *Fāngyán*, 1979(4):306-8.

118. Yip, Moira (1980a). *Tonal phonology of Chinese*. Doctoral Dissertation, MIT.

119. Yip, Moira (1980b). The metrical structure of regulated verse. *Journal of Chinese Linguistics*, 8:107-125.

120. Yip, Moira (1988) The obligatory contour principle and phonological rules: a loss of identity. *Linguistic Inquiry*, 19:65-100.

121. Yip, Moira (1989a). Tone contours as melodic units: evidence from Wuxi. In Marjorie Chan and Thomas Ernst, (eds.), *Proceedings of the Third Ohio State University Conference on Chinese Linguistics*, 13-14 May, 1988, pp.37-53. Reproduced by Indiana University Linguistics Club.

122. Yip, Moira (1989b). Contour tones. *Phonology*, 6:149-174.

123. Yuan, Jia-hua (1960). *Hànyǔ Fāngyán Gàiyào*, [Introduction to Chinese Dialects]. Beijing: Wénzì Gǎigé Chūbǎnshè.

124. Zadoenko, T. P. (1958). Hànyǔ ruòdú yīnjié hé qīngshēng de shíyàn yánjīu, [An experiment on the weak-stressed syllables and the neutral tone in Chinese]. *Zhōngguó Yǔwén*, 78:581-587.

125. Zee, Eric (1991). Tone feature system and Shanghai tone. Paper presented at *the Third North American Conference on Chinese Linguistics*, May 3 - 5, 1991. Cornell University.

126. Zhang, H. (1988). *A syntactic or prosodic domain? - On tone sandhi in Chongming*. Manuscript, University of California at San Diego.

127. Zhang, Zheng-sheng (1988). *Tone and tone sandhi in Chinese*. Doctoral Dissertation, Ohio State University.

128. Zhang, Zheng-sheng (1989). On the notion of 'word tone'. In Marjorie Chan and Thomas Ernst, (eds.), *Proceedings of the Third Ohio State University Conference on Chinese Linguistics*, 13-14 May, 1988, pp.5[illegible]77. Reproduced by Indiana University Linguistics Club.

129. Zhang, Zhen-xing. (1982a). Tone sandhi in Zhangping-Yongfu. *Fāngyán*, 1982(3):175-96.

130. Zhang, Zhen-xing. (1982b). Lexical tones in Zhangping-Yongfu. *Fāngyán*, 1982(4):264-5.

131. Zwicky, Arnold M. (1985). Rules of allomorphy and phonology-syntax interactions. *Journal of Linguistics*, 21:431-6.