

THE TONAL PHONOLOGY OF CHINESE

by

MOIRA JEAN(WINSLAND)YIP

B.A. Cambridge University
(1970)

M.A. Cambridge University
(1973)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January 1980

© Moira Jean Winsland Yip 1980

The author hereby grants to MIT permission to reproduce and to
distribute copies of this thesis document in whole or in part.

Signature of Author
Department of Linguistics and Philosophy
January 11, 1980

Certified by
Thesis Supervisor

Accepted by
Chairman, Departmental Committee

HUM.
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

JAN 11 1980

LIBRARIES

THE TONAL PHONOLOGY OF CHINESE

by

MOIRA JEAN WINSLAND YIP

Submitted to the Department of Linguistics and Philosophy on
11 January 1980 in partial fulfillment of the requirements
for the degree of Doctor of Philosophy.

ABSTRACT

This study sets out to construct a fully articulated theory of the phonology of tone, encompassing both the formal properties of tonal rules, and a feature system for tone. To counter-balance the African bias of earlier work, data are drawn mainly from Chinese and other East Asian languages. The first portion contains evidence that tone is autosegmental in Chinese languages as it is in African languages: the phenomena of stability, floating tones and toneless morphemes are shown to exist in Mandarin, Shanghai, Cantonese and Amoy. This is followed by a demonstration of the close relationship between tone and other prosodic phenomena such as **stress** and metrics.

The second portion constructs a binary feature system for tone based on two hierarchically organized features, each of which is autosegmental. Historical, typological and synchronic evidence are used to show that this system solves several long-standing problems in the study of tone.

The final portion includes a detailed analysis of the tonology of five languages -- Mandarin, Cantonese, Shanghai, Amoy and Fuzhou -- within the framework developed in the earlier sections.

Thesis Supervisor: Morris Halle

Title: Ferrari P. Ward Professor of Modern Languages and Linguistics

ACKNOWLEDGEMENTS

There is never any way to thank all the people whose physical or spiritual presence make a thesis possible, but a few can be singled out. Most immediately, I would like to thank the members of my committee, Paul Kiparsky and Kenneth Hale, and especially my adviser Morris Halle. They patiently submitted to a weekly exposure to fresh torrents of unfamiliar and confusing data, and helped me gradually make sense of it all and eventually emerge with some kind of consistent theory about what was going on. Morris Halle in particular could always be counted on to put his finger on the fatal flaw in my arguments, and sometimes more encouragingly to find the key to the problem that had eluded me in many hours of work. From the first day in 'baby phonology' when he made us learn the formant frequencies of the English vowels by heart, to the day when I handed him the last chapter of this thesis, most of my phonology has been learnt from him. The formant frequencies have long since been forgotten, but the phonology has, I hope, made a more lasting impression.

There are many other people who have helped clarify my ideas while I was working on tone, and in particular I'd like to thank Huang Zheng-de, Nick Clements, Mary Clark, and Matthew Chen.

Even with the linguistic support of all these people I could never have finished this thesis -- or indeed even started to study linguistics -- without the support and encouragement of my husband, George Yip. Although he considers it perverse to wish to compete in a declining market like linguistics, he has always helped in any way he

can (bar acting as unpaid informant), and without his constant presence I would have given up long ago.

Lastly (or rather firstly), I'd like to thank my parents, who silently accepted my career changes in ever less intelligible directions from maths to public relations to linguistics. From the beginning they gave me every opportunity to follow my own interests, and I will always be grateful to them.

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT	2
ACKNOWLEDGEMENTS	3
INTRODUCTION	8
0.1 The Nature of the Problem	10
0.2 Approaches to the Problem	16
0.3 The Choice of a Suprasegmental Theory	20
0.4 Notation, Romanization, and Other Trivia	30
0.5 Tone in Chinese: The Autosegmental Evidence	35
0.6 A Feature System for Tone	44
0.7 Some Formal Properties of the System	51
CHAPTER 1. ARGUMENTS FOR AN AUTOSEGMENTAL ANALYSIS OF CHINESE LANGUAGES	60
1.1 Segmentless Morphemes	60
1.1.1 Cantonese Changed Tone	60
1.1.2 Mandarin Reduplication	66
1.1.3 Amoy Triple Reduplication	75
1.2 Toneless Morphemes	79
1.2.1 Mandarin Suffixes	79
1.2.2 Southern Min Neutral Tone	83
1.3 Deletion of Segments Only	87
1.3.1 Cantonese Changed Tone	87
1.3.2 Amoy Contraction	91
1.4 Deletion of Tone Only	94
1.4.1 Mandarin Neutral Tone	94
1.4.2 Shanghai Word Tone	97
1.4.3 Amoy <u>a⁵³</u> Suffix	99
1.5 Spreading	102
1.6 Segmental Tone in Zahao?	104

1.7	A Dynamic Tone Theory	116
1.7.1	A Brief Summary of Dynamic Tone Theory	116
1.7.2	Problems Stemming from the Choice of Dynamic Primitives	126
1.7.3	Problems Stemming from the Positioning of Tone Markers Among the Segmentals	135
	NOTES TO CHAPTER 1	142
CHAPTER 2.	INTERACTION OF TONE AND OTHER PROSODIES	144
2.1	Tone and Stress	144
2.1.1	Mandarin	144
2.1.2	Amoy	157
2.1.3	Shanghai	163
2.2	Traditional Chinese Verse	166
2.2.1	A Metrical Analysis	166
2.2.2	Neutral Tone in Min	175
2.2.3	Deviations from the Canon: The Free Positions	179
2.2.4	Tone Assignment	186
	NOTES TO CHAPTER 2	193
CHAPTER 3.	AN INTRODUCTION TO THE CONCEPT OF REGISTER	195
3.1	A Summary	195
3.2	Typological Considerations	198
3.2.1	Five-Tone Languages	202
3.2.2	Three Contour Tone Languages	204
3.2.3	Downstep and Other Matters	207
3.3	Historical Evidence	211
3.3.1	Why Other Systems Fail	211
3.3.2	Voicing -- Necessary But Not Sufficient Condition for Register Development	216
3.3.3	Mandarin and Cantonese Tonal Splits	223
3.4	The Physiology of Pitch Control	242
	NOTES TO CHAPTER 3	246

CHAPTER 4.	THE AUTOSEGMENTAL NATURE OF REGISTER	250
4.1	Partial Morphemes	250
4.1.1	Morphemes Lacking One Level Only	251
4.1.2	Morphemes Lacking Two Levels	255
4.2	Partial Deletion	261
4.2.1	Deletion on One Level Only	261
4.2.2	Deletion on More Than One Level	267
4.3	Spreading	272
4.4	The Domain of Register and Tone	273
	NOTES TO CHAPTER 4	282
CHAPTER 5.	SUMMARIES OF THE TONAL PHONOLOGY OF VARIOUS LANGUAGES	283
5.1	Mandarin	283
5.2	Shanghai	300
5.3	Amoy	319
5.4	Fuzhou	337
5.5	Cantonese	356
	NOTES TO CHAPTER 5	359
REFERENCES		361

INTRODUCTION

A fully articulated theory of tonal phonology must consist of two largely independent but consistent parts. On the one hand it must include a formal statement of the properties of tonal rules, and on the other hand it must include a characterization of the distinctive features of tone that form the raw material on which the rules act. To be satisfactory, such a theory must correctly predict the kinds of tonal systems that can and do occur, as opposed to those that cannot and do not occur. In building such a theory it is therefore necessary to take into account the tonal phonology of as wide a variety of tonal languages as possible, and the addition of any new language to the data base provides a fresh test of any pre-existing theory, which may then require revision in the light of the new data. If, on the other hand, the tonal phonology of the new language is such as had been predicted by the old theory, that theory is strongly supported.

Now it is noteworthy that, with a few exceptions such as Woo (1969), Wang (1967) and Walton (1976), most of the recent theoretical work on tone has been based largely or entirely on data from African languages. This thesis is an attempt to address an entirely new set of problems posed by tonal phenomena in Chinese-type tone languages. It is to be hoped that the Mexican tone languages, to cite another large and under-researched group, will receive similar attention in the future. The structure of the thesis is as follows. First I will take the theory of Autosegmental Phonology essentially as developed by Goldsmith (1976) based on data from African languages, and put it to the test of dealing also with very different data from Chinese

languages. We shall see that the essential predictions of the theory are borne out, and conclude that tone in Chinese languages, too, is autosegmental. This will therefore provide the first element of the complete theory: a formal statement of the properties of tonal rules that works in a variety of very different tone languages. Chapters one and two are concerned with this question.

In chapter three I develop the second element of the theory: a feature system. Starting from typological considerations, and moving on to historical arguments, I propose a hierarchical feature system that separates out the general pitch level, which is given by a Register feature, from the tonal melody, which is given by a sequence of Tone features. Assuming the results of chapter one -- that tone in general is autosegmental -- I proceed in chapter four to show that a more articulated theory in which each of the two hierarchical features is independently autosegmental provides an elegant account of a number of tonal phenomena in various Chinese languages.

Finally, in chapter five I give a quite complete account of the tonal phonology of several languages in order to show that complete systems can be accounted for. Some of the phenomena will have already been discussed in earlier sections, but here I will show how they work as parts of a larger whole.

In the remainder of this chapter I discuss earlier work in the phonology of tone, and summarize the major arguments in favor of an autosegmental approach. This is followed by a brief exposition of the major findings of this thesis, cross-referenced to the supporting arguments in the body of the text. The summary is intended to provide

a quite complete picture of the framework developed here, and those readers with a particular interest in Chinese may then refer to the relevant sections of the text for a fuller treatment.

0.1 The Nature of the Problem

The two separate questions that must be answered in developing a theory of tone are what is the relationship of tone to the rest of the phonology, and what are the features of tone. Although these questions are logically independent, it turns out not surprisingly that the answer to either has implications for the answer to the other, given the facts of natural language. It makes sense, therefore, to look first at some of those facts and summarize the essential properties of tone that must be explained by a satisfactory theory. Once this is done we can then assess the various possible theories.

Tone behaves very differently from other components of the morpheme. There are six particularly noticeable characteristics of tone which I will now discuss. Each of them has bearing on the answers to the above questions; the first two properties bear on the relationship of tone to the rest of the phonology: specifically, is tone, like vowel height, a feature of a segment (or perhaps a syllable) or is it separate and suprasegmental? The remaining four properties bear on the choice of features for tone, and especially whether contour tones are to be treated as units, or analyzed into sequences of level tones. (The first property in this group, the existence of contour tones on short vowels, actually bears on both issues.)

a) Stability: In many languages changes in the segmental level leave the tones unaffected. This can be seen particularly clearly in

word games which interchange segmental material, but leave the tones behind. For example, Thai has one type of word game which effects the following changes in which the rhyme moves but the tone remains behind:

klúà̀y hḿ́m → klḿ́m huá̀y 'banana'
 tén rām̄ → tám rēn̄ 'dance'

English speech errors provide a closely related example. If in a speech error the two parts of a phrase or compound are interchanged, the stress pattern is unaffected. English stress consists of a falling tone with the beginning of the fall always realized on the stressed syllable. If tone were a feature on the syllable, it should move with it, but instead it stays put:

hámmer and síckle → síckle and hámmer
 bírdseed → seédbírd

(This argument is essentially due to Becker (1979). For further examples of the stability of stress patterns in speech errors and malapropisms see Fromkin (1971) and Fay and Cutler (1977).)

Any treatment of tone as a feature of the syllable predicts that when a syllable is lost the tone should be lost too, and when a syllable is moved the tone should move too. Neither of these predictions is correct. Similar problems arise if tone is treated as a segmented feature. In Lomongo (Lovins 1971) under certain circumstances consonants may delete and trigger the elision of the first of two vowels. Far from deleting with the vowel, the tone

remains and is realized on the second vowel:

bàlóngó bǎkáé → bàlóngǎkáé 'his book'
 bǒmǒ bòtámá → bǒmǒtámá 'another tree'

This type of phenomenon remains unexplained in a segmentally based approach.

b) Floating tones and tense melodies: It has been clearly shown that some languages have floating tones and tonal melodies. These are morphemes that have no phonemes, but do have tone. This tone is demonstrably present, and may actually surface on the nearest morpheme, or it may show up by virtue of its secondary effects on neighboring tones. For example, Goldsmith (1976b: 61) shows that Igbo has a subordinate clause marker consisting solely of a H tone. When this precedes a cliticized low-toned pronoun the pronoun acquires this floating tone as well, producing a falling tone on the surface as a result of the sequence HL:

 / ò gbùò éghú → ô gbùò éghú 'lest he kill the leopard'
 SUB he kill leopard

For an Asian example I turn to Burmese. (The data are from Bernot 1979. The interpretation is mine.) It is apparently the case that nominal modifiers in Burmese are marked by tone one (written with an acute accent following the syllable as in -tɛ́). This tone shows up on the clause-final particle if present, or on a pronoun when this is used alone as a possessive. By contrast the clause-final particle in non-modifying position bears tone two (unmarked), and so do pronouns

in isolation.

seN `θi ne Tɛ 'she is covered in diamonds'
 seN `θi ne Tɛ' θə̀mi 'a girl covered in diamonds'
 (diamond produce AUX PARTICLE girl)
 ŋa' ʔə̀pyi? 'my fault' (ŋa 'I., ne')

It seems from these facts that tone one is itself a marker of nominal modification, and it may show up on any adjacent morpheme.

Facts of this kind are a problem in a segmentally or syllabically based feature system for the following reason. These morphemes consist of tone without a syllable (or segment) but if tone is a feature of the syllable (or segment) there is no way of describing such morphemes. These first two problems are problems for any claim that tone is a feature of the syllable or segment. The remaining problems result from any attempt to claim that contour features are units.

c) Contour tones on short syllables: The existence of contour tones on short vowels is superficially an argument in favor of a unitary contour approach, since any segment or syllable may bear a feature [+ Rising] (for example). However, some languages, such as modern Lithuanian, only permit contour tones on longer syllables. Specifically, contour tones are only found on syllables containing sonorant sequences in the rhyme, either VV sequences or vowel plus nasal or liquid. Such a restriction is decidedly unexpected in a contour feature framework, within which it would be just as possible to restrict one of the level tones to long syllables. Later work, which decomposes contour tones into sequences of level tones (so that

a rising tone is made up of the sequence LH) allows a simple statement of this restriction. This will be discussed in greater detail later.

d) Identical behaviour of contour tones and sequences of level tones:

It has been shown in a number of cases that taking contour tones to be composed of sequences of level tones allows much simpler statements of such things as the distribution of tonal melodies in languages. For example, consider the tone classes of Mende words (Leben 1973). Short vowels can have high, low, rising, falling, or rising-falling tones. However, it is not the case that in a two-syllabled word either vowel can have all of these possibilities (if it were, there would be a total of 25 tonal patterns in bisyllables). In fact only five patterns exist with the vowels being respectively H + H, L + L, L + H, H + L, or L + Falling. In three-syllabled words there are also only five patterns. All these facts can be accounted for by assuming that the tones are in fact the result of only five possible morpheme melodies which show up on the vowels in a simple, predictable way determined by the number of syllables in the word. (The first tone shows up on the first vowel and so on, left-to-right, with any left-over tones showing up on the last vowel.) The melodies are H, L, LH, HL, LHL:

H	pé é	kó	
L	bè è	kɛ̀	
LH	ní á	mbǎ	
HL	ké nyà	mbú	
LHL	ní kí lí	nyà há	mbǎ

In other words, the contour tones found on single syllables are the

precise equivalent of the sequences of level tones found on longer morphemes. There is no obvious way of capturing these facts if the contour tones are portrayed by means of contour features.

e) Reference to end-points of tones: If contour tones are expressed by unitary features there is no way to refer to the end-points of tones. Furthermore, it can be shown that contour tones often behave exactly like their level tones that match their end-points in conditioning or undergoing tone rules; if they are indeed composed of level tones this is unsurprising, but if they are separate entities such behaviour is hard to explain. In Central Chin (Eugenie Henderson, personal communication) there is a whistling language used to communicate over long distances. But what is whistled is not the entire tone of a word but just its end-point. The reverse situation can be found in the historical development of Paangkua (Löffler 1979), in which the four tones of Lushai have apparently been reduced to two by a process of simplification to the starting point of the tone, so that Lushai high and falling tones are both high in Paangkua, while the low and rising tones are both low. Another simple example comes from Sui rhymes (Li 1949). In Sui verse each tone may rhyme not only with itself but also with one other tone, as follows:

- | | | | | |
|----------------|---|---|-----------------|---|
| 1. low | ┘ | = | 2. low falling | ┘ |
| 3. mid | ┘ | = | 4. high falling | ┘ |
| 5. high rising | ┘ | = | 6. high level | ┘ |

It can easily be seen that what determines the rhyme is the end-point

of the tone, with each pair ending on low, mid and high levels respectively.


f) Over-enrichment of the theory: If, as suggested by (c-e) above, some contour tones must be represented by sequences of level tones, it would obviously be preferable to restrict the theory so that all contour tones are so represented. That some contour tones are sequences is quite clear. For example, the Igbo floating tone discussed in (b) produces falling tones as a result of the addition of a high tone to an underlying low tone. Obviously this fall is therefore to be represented as HL. Admitting the additional possibility for language of [+ Falling] then predicts a possible contrast between HL and [+ Falling] that does not occur.

I have set out the six problems in (a-f) in some detail because I will refer to them again and again in discussing attempts at building a theory of tone.

0.2 Approaches to the Problem

I began the previous section by pointing out that there were two separate questions that had to be answered in developing a theory of tone: what is the relationship of tone to the rest of the phonology, and what are the features of tone. There are essentially two possible positions one can take on each of these questions. The answer to the first may be either that tone is a feature on segments (or syllables) like any other feature, or that tone is quite separate, a suprasegmental, prosodic phenomenon. The answer to the second may be either that tones are units, so that contour tones have a single

feature, such as [+ Rising], or that all tonemes are level, and contour tones are composed of sequences of level tonemes. Combining these answers produces four different theories of tone, as follows:

Features	Relationship of Tone to Phonology	
	Segmental/Syllabic	Suprasegmental
Unitary Contours	A. Wang	B. 
Sequences of Level Tones	C. Woo	D. Williams Leben Goldsmith

The names on the above chart refer to various theories of tone within the framework of generative phonology. Of course there have also been many earlier, non-generative discussions of tone, including that of Daniel Jones, which belongs in box B on the chart, and some of these will be referred to below. First I will summarize the essential properties of each of the theories referred to on the chart, and then I will show that the choice between them is largely determined by the properties of tone laid out in the previous section.

In 1967 William Wang proposed a feature system for tone which had the following properties: (i) Tones were features on syllables, rather than on segments; (ii) All tones, including contour tones, were units. The feature system thus included not only specification of a number of level tones (five, in fact) but also features referring to the shape of the tone, such as [+ Contour], [+ Rising], [+ Falling], [+ Convex].

Wang's system had certain advantages. Taking tone to be a property of the syllable explained naturally why tone rules do not normally make reference to segmental features, and also fitted the phonetic facts rather well (i.e., that tone is realized over the entire voiced portion of the syllable). Treating contour tones as units had the advantage that contour tones on single, short vowels could be stated just as easily as level tones on such vowels: we shall see later that this is not true of some later theories. However, it also raised a number of rather fundamental problems, as we shall see later.

In 1969 Nancy Woo proposed an entirely different approach in her thesis, Prosody and Phonology. The two main characterizations of her proposal were as follows:

(i) Tone was a feature on segments (sonorant segments in the rhyme, to be precise).

(ii) Contour tones were composed of sequences of level tones.

With a feature system that defined five levels (given in full in section 3.3.1), and the idea of decomposing contour tones into their end-points (and sometimes also a mid-point, in the case of convex or concave tones) she was able to depict all of the rather complex tones found in Asian tone languages. Notice however that given the two claims in (i) and (ii) it follows automatically that any contour tone, requiring two or more level tonemes as its specification, required two or more sonorant segments. The immediate consequence of the theory was then a claim that contour tones could not occur on single short vowels, and indeed Woo claimed exactly this. Although it is true that

the tones found on short vowels are much more restricted than those found on longer syllables, with contour tones being relatively rare, it is quite clear that they do in fact exist. In Cantonese, for example, under certain circumstances mono-moraic syllables may bear a rising tone:

玉 yuk 'jade'

The final voiceless stop cannot possibly bear tone, so there is no way to represent the rising tone, which requires a sequence of tonal features [+ low pitch] [+ high pitch] on a single sonorant segment [u]. Notice in support of Woo that languages like Lithuanian which do limit contour tones to sonorant sequences are quite common; however, the existence of languages like Cantonese that lack this restriction forces us to relinquish one or both of Woo's claims -- the segmental nature of tone, or the compositional nature of contour tones.

Williams (1971), Leben (1973) and Goldsmith (1976), the three theories in the lower right-hand box, D, on the chart, are closely related, though distinct in certain respects. All three share the following properties:

- (i) Tone is a suprasegmental, and a separate part of the lexical entry for each morpheme.
- (ii) Contour tones are composed of sequences of level tones.

Later I will discuss the differences between these three, but for the moment their shared properties are more important. Notice that these approaches differ from Woo in relinquishing the claim that tone is

segmental. Once this is done the problem of accounting for contour tones on short vowels ceases to exist.

We are now in a position to relate these theories to the properties of tone, and see where each fails or succeeds. Recall that the properties of tone fell into two groups. The first group, included (a) stability and (b) floating tones. These two properties can only be satisfactorily explained if tone is suprasegmental, not segmental (or syllabic). These properties alone therefore exclude Wang's and Woo's theories. The last three properties, (d) the equivalence of a contour and a sequence of level tones, (e) reference to the end-points of tones and (f) over-enrichment of the theory, can only be explained if contour tones are composed of sequences of level tones. Wang's theory is thus excluded on these grounds too, and so is the type of theory that belongs in the upper right-hand box, B, which claims tone is suprasegmental, but that it has unitary contour features. The last property, the existence of contour tones on short vowels, implies one of two alternatives: either tone is suprasegmental, or if it is segmental there must be unitary contour features. Since the second alternative is excluded by properties (d-f) we are again led to the conclusion that tone is suprasegmental.

It is therefore quite clear that only a theory in which tone is suprasegmental, and contour tones are made up of sequences of level tones, will satisfactorily account for the observed properties of tone.

0.3 The Choice of a Suprasegmental Theory

The view that tone is suprasegmental is far from being a novel

idea, but is in fact a return to the earliest ideas about tone. Traditional Chinese scholars considered the tone to be a property of the rhyme (the syllable minus its initial consonant(s)), and this shows up quite clearly in the fān-qie 反切 spelling system, in which a character was 'spelt out' by giving two other characters, the first of which shared the same initial consonant and the second of which shared the same final. The tone of the second character was always the same as that of the target character, showing clearly that it was perceived as a property of the rhyme, and never of the initial consonant (or even of the syllable as a whole, in which case one might expect a random distribution, or a requirement that both of the characters used in fān-qie have the same tone). As an example of the spelling system in practice:

† qiān = 蒼 c (āng) + 先 (x) iān

(The difference in initial consonant is the result of palatalization before the high front medial glide.)

In the West, Zellig Harris's long components (1944), and before that the prosodic school of Firth influenced by the work of students of Daniel Jones all took the view that tone is a suprasegmental. Essentially this meant that tone was not tied to one particular segment, but rather was a property of some larger unit that happened to be realized eventually on individual vowels and sonorants. With the rise of generative phonology in the 1960's, in particular the framework developed in Sound Pattern of English (Chomsky & Halle, 1968, henceforth SPE), the emphasis shifted to segmental phonology, and it

was assumed that segments were packages of distinctive features, and that all features formed part of some such package. Tone was no exception: for languages like Chinese with lexical tone they proposed (p. 377):

'to mark the prosodic features for each vowel in the lexicon.'

Nancy Woo's thesis was written very much in this framework, but it subsequently became clear that the attempt to fit tone into a segmental mould was a failure, for the reasons discussed above. So attention shifted back to the earlier assumption that tone belonged not with segmental features like coronality or labiality, but with suprasegmental features like intonation.

In 1971 Williams wrote a paper called "Underlying Tone in Margi and Igbo" (published in 1976) in which he showed that tone could not be associated underlyingly with segments or even with syllables, but must rather be a property of the morpheme. This was followed in 1973 by Will Leben's thesis Suprasegmental Phonology, in which he proposed a theory with the following properties:

- (i) Tone is a separate part of the lexical entry for each morpheme.
- (ii) Tone is mapped onto the vowels of the morpheme from left to right.
- (iii) Contour tones are the result of mapping more than one tone onto a single vowel: that is they are sequences of level tones.

This formalization of the suprasegmental position was a giant step towards solving the remaining problems encountered by Wang. Like

Woo, Leben decomposed contour tones into sequences of level tones and thereby solved the three problems (c-f) that involve necessary reference to the component tones. But he also solved the first two problems that had stemmed from the segmental analysis of Woo (and the syllabic one of Wang). If the tone is a separate part of the lexical entry, rules that apply before the mapping merges those parts not only can, but must leave tone unaffected by segmental rules, and vice-versa. A rule eliding the first of two vowels, as in Lomongo, says nothing about the tonal component, so when the vowel is lost 'its' tone will remain behind and eventually be mapped onto some remaining vowel. Further, if tone is a separate part of the lexical entry it is logically possible for a lexical entry to lack just one of its parts. If it lacks tone it is a toneless morpheme, and if it lacks segments it is a floating tone. For example, McCawley (1970) showed that in Tiv the tonal patterns of verbs can best be explained as the result of a combination of the basic tone of the verb with the characteristic tone of the tense or aspect. The aspectual suffixes are therefore composed solely of tone, a natural possibility in this theory, but a problem in a segmental approach. (The most common solution within a segmental framework is to allow representations like [- segment, + high tone].) It should be clear, then, that problems (a-b), stability and floating tones, are far from being problems in Leben's framework; on the contrary these phenomena are actually predicted by the theory, and strongly support it.

The one remaining problem is (c) the existence of contour tones on short vowels. This is still a problem because of the mapping

procedure, which will result in two features on a single segment in such cases. For the solution we turn to Goldsmith (1976a) (although Williams (1971) probably intended something similar). In his thesis Autosegmental Phonology Goldsmith took Leben's work one stage further and argued that tone was not only a separate part of the lexical entry, but that it stayed independent throughout the derivation and even in the phonetic representation, and was linked to the phonemic level by association lines rather than by ultimate mapping. This is the essential claim of autosegmental phonology as opposed to other suprasegmental approaches, as it solves the problem of describing contour tones on short vowels, since these now involve the association of two tones with one vowel, rather than an actual mapping.

Notice a further consequence of the autosegmental approach. I have argued that some contour tones must be analyzed as sequences of level tones, and that it would obviously be preferable if all contour tones could be so analyzed. Suppose for the moment that this were not the case. Then in addition to level tonemes there would be contour tonemes, and presumably these would be associated by the same rules. In that case multiple associations of contour tones would arise just as easily as multiple associations of level tones, and some very complex tones should result (for example, [+ rising] [+ rising] would presumably be phonetically rise-fall-rise and so forth). Tones of this complexity are rare or unknown, whereas contour tones resulting from sequences of level tones are very common.

Let me digress at this point and add some final remarks about this question of level versus contour features. One might of course

argue that languages are of two types: those which use level tones (and are presumably autosegmental) and those which use contour tones (perhaps segmental). Pike (1948) made essentially this distinction between level pitch systems and gliding pitch systems. In a level pitch system there are a restricted number of contrastive levels, called registers. (Note that this is an entirely different use of the term from that adopted here in chapters 3 onwards.) Any contour tones begin on one register and end on another: they can therefore be viewed as transitions from one level to another, and the actual contour is irrelevant. Most African languages and many Mexican languages, like Mixteco, are apparently of this type. In a gliding pitch system, on the other hand, "the beginning and ending points of the glides...cannot be identified with level tonemes in the same system" (p. 8). He also suggests that such tonemes are limited to one per syllable. Mandarin would be an example of such a language, with a tonal inventory consisting of one rising, one falling, one level, and one falling-rising tone. Pike goes on to admit that most languages have characteristics of both systems and cannot easily be dealt with in one or the other way alone.

While I cannot prove the absence of languages of the second type -- segmental tone and contour features -- I have failed to find any such languages.

We will show that a contour tone analysis of changed tone in Cantonese (the type of language often analyzed as having unitary contour tones) fails to offer an elegant solution to what is, after all, quite a simple set of facts (see 1.1.1). We will conclude that

there is no evidence whatsoever to force the addition of contour tones to the system, and that in fact they only complicate the analysis.

Recently there has been some discussion of the facts of a Tai dialect of Lue. Gandour and Fromkin (1978) have argued that the only tonal alternation in this language cannot be simply analyzed given only level features, and that in fact what is going on is a type of contour dissimilation: before a level tone the contour variant [13] occurs, and before a contour tone the level variant [11] occurs. They therefore propose a rule which changes the tone to [- Contour] before [+ Contour]. Maddieson (1979) has argued that another analysis is available which does not rely on contour features, and offers an equally natural characterization of the facts. The interested reader is referred to the original articles: let us simply quote Anderson (1978), who says (p. 159):

"...its (i.e., Tai Lue's) very isolation in the tone systems of the world, together perhaps with the fact that it is not obviously assimilatory or otherwise phonetically motivated, makes it difficult to support such features (i.e., contour features) on the basis of this example alone."

(The work from which this quote comes is an extremely useful survey of the major arguments over level versus contour features, and other matters.)

Another recent approach to the question is that of Mary Clark (1978) who has proposed what she calls a dynamic tone analysis which is essentially closer to a unitary contour approach than a level tone approach. This will be discussed separately in section 1.7.

Returning to my main theme, Goldsmith's Autosegmental Phonology

accounts for all the observed characteristics of tone discussed so far, and in what follows I will be concerned with testing its predictions against data from Chinese, and with further refining the theory where necessary. Let me therefore pause and briefly lay out the terminology and formalism of the theory for those unfamiliar with it. Other readers may skip to section 0.4.

Lexical entries consist of one or more separate tiers, one for the phonemes and another for the tones in the typical case. Each tier is segmental in the sense that it consists of a sequence of discrete units, but I will usually refer to the phonemic tier as the segmental tier for the lack of a satisfactory alternative. The tiers are normally entirely unassociated in the lexical entry (although in some languages or in some irregular lexical items association lines may form part of the lexical entry). The association lines are then inserted at the very beginning of the derivation by a combination of language specific rules and a universal convention known as the Well-Formedness Condition (henceforth WFC). This condition applies at every stage of a derivation to fix up the output of any rule, and is stated by Goldsmith as follows:

Well-Formedness Condition (WFC)

1. Each vowel must be associated with (at least) one toneme.
2. Each toneme must be associated with (at least) one vowel.
3. No association lines may cross.

Goldsmith gives the WFC in terms of associations between tonemes and vowels, but the tone-bearing unit might in fact be the sonorant,

rhyme or syllable, and probably differs from language to language. In Chinese the rhyme or the syllable is probably the relevant unit.

The most common kind of language specific association rule associates the first tone with the first vowel, then moves rightwards across the word, associating one-to-one. The WFC applies to any left-over vowels or tones. For example, consider the Mende words discussed earlier. When there are an equal number of tones and vowels the situation is straightforward, and handled entirely by the language specific rules:

ko	nika	nikili
⋮	⋮ ⋮	⋮ ⋮ ⋮
H	L H	L H L

(The broken lines are used to show lines inserted by the rule under discussion. Solid lines are pre-existing ones.)

When there are extra vowels the language specific rule only does half the work:

pele
H

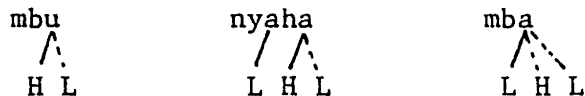
The WFC provision (1) will then fix this up as follows:

pele
/
H

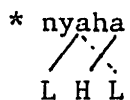
When there are extra tones the language specific rule gives:

mbu	nyaha	mba
/	/ /	/
H L	L H L	L H L

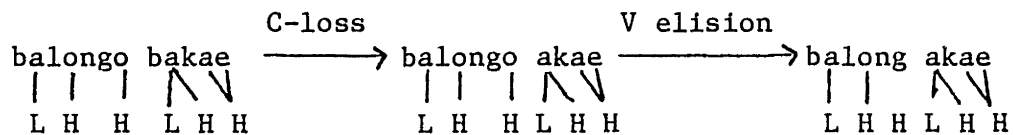
The WFC (2) then requires the insertion of the remaining lines:



Note that provision (3) blocks any other association of the spare tone, such as:



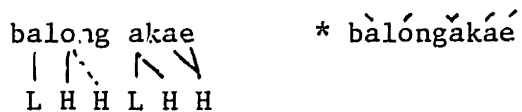
The WFC operates after any phonological rule as well. For example, the derivation of the Lomongo forms with vowel elision proceed as follows:



But this now violates the WFC since there is an unattached H tone, and will therefore be fixed up as follows:



The alert reader may have noticed that there is an alternative way to satisfy the WFC in this last case:



But this runs counter to the facts. In other words, the WFC as it

stands fails to fully specify the necessary action in some cases. This has been discussed by Clements and Ford (1979), who also argue that condition (2) requiring all tones to be associated is too strong, and that some tones may remain unassociated throughout the derivation unless associated by a language-specific rule. They therefore develop a more precise set of conventions which removes the ambiguities from the WFC. For the time being I will stick to Goldsmith's original version of the WFC, but later I will discuss Clements and Ford's claims in some detail in connection with certain facts in Chinese languages, and conclude that they are right in suggesting that condition (2) is too strong.

0.4 Notation, Romanization, and Other Trivia

Since the remainder of this thesis is concerned largely with Chinese languages it behooves me to give some background to the romanization, tonal diacritics, and other conventions of Sinologists. For ease of reference all Mandarin forms will be given in the pinyin romanization now given the official imprimatur of the New York Times (which does not, however, bother with the tone marks, thereby leading to some misunderstandings). The tones are shown by diacritics above the vowel, as follows:

tā	'he, him'	First tone.	High level pitch.
cháng	'long'	Second tone.	Rising tone.
lǎo	'old'	Third tone.	Low, with a rise in isolation
dà	'big'	Fourth tone.	Falling tone.

The segmental phonology is largely irrelevant to our concerns, so we will not include a table of IPA equivalents to the pinyin system.





Note only that Mandarin (like most Chinese languages) has no voiced initial obstruents: the orthographic distinction between t and d, etc., is phonetically a distinction between aspirated and unaspirated voiceless stops.

The other Chinese languages are not as familiar, and in most cases there is no single agreed system of romanization. In general we will use the same orthography as the cited source, supplemented where necessary with explanations. Two types of representation of tone are common to most sources, and are summarized here. Firstly, many writers use the so-called tone letters of Y-R Chao (1930). A vertical stave represents the natural pitch range of the speaker's voice, and a line is drawn to one side of this stave to show the level and contour of the tone (read from left to right). For example:

tā	High level	┌ (or ┐)
cháng	Rising	↗ (or ↘)
lǎo	Low, plus final rise	↘ (or ↗)
dà	Falling	↘ (or ↗)

These are the same four words used to exemplify Mandarin diacritics on the preceding page.

Sometimes this stave is numbered, in which case it is numbered from a low of 1 to a high of 5, with equal divisions in between. The tones can then be referred to numerically, without the stave being actually drawn:

tā	5		55
	1		
cháng	5		35
	1		
lǎo	5		214
	1		
dà	5		41
	1		

For ease of typing these digits are often superscribed after a word (or occasionally inserted at the same level): ta⁵⁵ or ta55.

A few general comments about Chinese languages. Under the general banner of Chinese are subsumed at least seven language groups, and each of these includes a myriad of sub-languages and dialects. It is not at all clear at what point the term language should cease to be used, and the term dialect take its place, since the terms are essentially political ones. Of the languages discussed here I think it is true to say that not one pair is mutually intelligible, and I have therefore mainly used the term language. That said, they do of course have a great deal in common. The syntax is surprisingly similar for such a vast area, and most of the differences that result in mutual unintelligibility come in the phonology. For example, three of the language groups have retained, to a greater or lesser extent, syllables with final obstruents (which always carry a restricted tonal inventory); the others have lost them, and now possess only

sonorant final syllables. Two groups, the Wu and the Xiang dialects, have retained voicing in their initial stops while the others have lost it. The tonal systems (as we shall see) have developed in totally different directions, to the extent that any reconstruction of a proto-system and its subsequent development looks almost hopeless. And the remainder of the syllable, the vowel nucleus and adjacent glides, has also drifted in a number of different directions. For example, my last name is written 葉 in all dialects, and means 'leaf'. In Cantonese (from which the romanization is taken) it is pronounced yip²². In Mandarin it is ye⁴¹, with loss of final stop and also vowel change. Or the word for 'day', written 日. In Mandarin this is ri⁴¹ (with an apical vowel and voiced fricative r); in Cantonese it is yit²². In Suzhou, a Wu dialect, it is ni?²³; and in Yangzhou, an Eastern Mandarin dialect, it is la?⁴. There has also of course been considerable change in vocabulary over the centuries, so that where Mandarin has tā 他 as the third person pronoun, Cantonese has koei²⁴ 佢, an entirely different character.

In one major respect the languages still agree: they are largely monosyllabic in the sense that the equivalence between morpheme and syllable is almost perfect. There are in Mandarin, and presumably in the other languages also, a few morphemes of two syllables, and a very few sub-syllabic ones (such as the retroflex suffix r, and floating H tone in Cantonese). Otherwise the morpheme = the syllable.

For anyone working on the Chinese languages there are certain indispensable sources of data, unfortunately published in Chinese only. These are Hànyǔ Fāngyán Gàiyào (HYFYGY), Hànyǔ Fāngyán Cíhuì

(HYFYCH) and Hànyǔ Fāngyán Zìhuì (HYFYZH) (the last is chiefly of use for historical reconstruction).

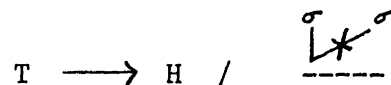
It is unfortunately a problem in many cases to decide exactly what the facts are. Very little instrumental work has been done on Chinese tones and the different field workers often disagree quite widely on their perceptions of the tones. Where there is such disagreement it is often unclear whether two different dialects of idiolects are being described, or whether the descriptions are of the same data, differently perceived. Sometimes the disagreements are unimportant. For example suppose one field worker claims that a language includes tones 13 and 24, while one other says it has 13 and 35; it seems reasonable to assume that the important fact is one on which they both agree: the language has two rising tones, one higher than the other, and the finer details are unimportant for present purposes. However, in many cases there is disagreement over the shape of a tone and not just its level. In Fuzhou, for example, the Shang tone is reported as falling 21 by T'ao (1930), falling 31 by HYFYCH, and level 22 by Chao (1933) and T. C. Wang (1969). In such cases I have taken as the right contour whichever has most support from the sources, or appears from the phonology to be the predicted form. Quite often tones which are reported as having a slight fall (such as 21) behave phonologically as level tones. This is the result of a quite general phenomenon in language called declination (Pierrehumbert 1979).

A few final remarks on the form of the tone rules. Since many tone rules delete or add association lines, some way of showing this

is necessary. I will adopt the following conventions: association lines to be added are drawn as broken lines, while association lines to be deleted have a superimposed X:



The X will also be used in the context of some rules to show that that association line must be absent in order for the rule to apply:



This rule will raise any tone that is associated with only one syllable to H.

0.5 Tone In Chinese: The Autosegmental Evidence

Although the arguments for autosegmental phonology are quite clear, in the past they have been drawn almost entirely from African languages. It is therefore important to show that similar arguments can be found in other language families, such as Chinese. In chapter 1, I give examples of the major predictions of autosegmental theory as borne out in Chinese languages. Because of the paucity of morphological and phonological alternations in Chinese languages it has not always been obvious that tone leads an independent existence to the extent that it does in the African languages studied by Leben and Goldsmith. Since there is little inflection, there are no tense melodies, for example. Syllables tend to stay discrete from each other, so there is little elision of vowels resulting in spare tones.

However, we will see that when such processes do take place the results are exactly what we expect if tone is autosegmental, and would be inexplicable in a segmental system.

The major predictions of autosegmental phonology are two. Firstly, the independence of the tiers in the lexical entry introduces the possibility of lexical entries which lack specification on one or other tier. The result is a floating tone (or a toneless morpheme). The second prediction is that commonly known as stability: segmental material may be deleted, but the tone(s) will be unaffected. A simple example from Cantonese will suffice to illustrate both (for a detailed exposition see 1.1.1 and 1.3.1).

Cantonese exhibits a set of morphologically conditioned tonal alternations commonly called changed tone. These are of two kinds. In the first kind the only difference between the two forms is the tone, whereas in the second kind the unchanged (basic) tone co-occurs with certain segmental material, but in the changed tone form the segmental material is missing. An example of the first type is the occurrence of changed tones on family names in familiar use following the prefixes A:⁴⁴ or lou²⁴ (the asterisk denotes changed tone):

$$\begin{array}{l} \text{A:}^{44} + \text{ts}\check{\text{e}}\text{n}^{21} \longrightarrow \text{A:}^{44} \text{ts}'\check{\text{e}}\text{n}^{*35} \quad \text{阿 陳} \\ \text{lou}^{24} + \text{lei}^{24} \longrightarrow \text{lou}^{24} \text{lei}^{*35} \quad \text{老 李} \end{array}$$

An example of the second type is adjectival reduplication, which has two forms:

hung²¹ a⁵ hung²¹ ~ hung^{*35} hung²¹ 'very red' 紅 0 紅

In every case of this type the missing segments have a high tone.

Now the actual shapes of the changed tone can be summarized as follows:

- (i) If the underlying tone begins on a high pitch, the changed tone is a high level tone.
- (ii) If the underlying tone begins at any other pitch, the changed tone is a high rising tone.

So an underlying 53 falling tone becomes *55, whereas an underlying 24, 44, or 21 becomes *35 (roughly).

It is clear that (i-ii) above can be reduced to the single statement:

The changed tone begins at the same point as the underlying tone, and ends on a high pitch.

This fact has a very simple explanation in autosegmental theory. The changed tone is obviously formed by the addition of a high tone to the end of the existing tones (followed by some simplification):

$$\begin{array}{l} 53 + 5 \longrightarrow 535 \longrightarrow 55 \\ 24 + 5 \longrightarrow 245 \longrightarrow 25 \end{array}$$

There are two sources for this high tone. In the first kind of case, exemplified by the family names, the source is obviously a floating tone. In the second kind of case, exemplified by the reduplicated adjectives, the high tone is left behind when the segmental material is deleted, and reassociates with the nearest segmental material.

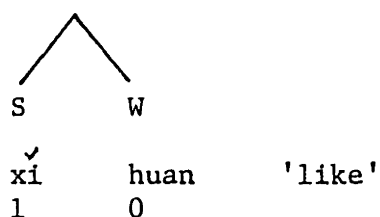
These are exactly the two basic predictions of autosegmental phonology: partial lexical entries, and stability. Neither can be dealt with easily in Wang's or Woo's frameworks.

After discussing a number of examples from various Chinese languages that show rather conclusively that tone is autosegmental in those languages, I continue to discuss two languages which have been reported to have segmental tone, Zhaoh and Thai, and conclude that an autosegmental analysis is at least as plausible in each case. Although there is no logical necessity for tone to be autosegmental in all languages (any more than nasality is) no clear example of segmental tone is known to me (Northern Tepehuan (Woo 1969: 94) is the most plausible case), and it is at least evident that the unmarked case is for tone to be autosegmental.

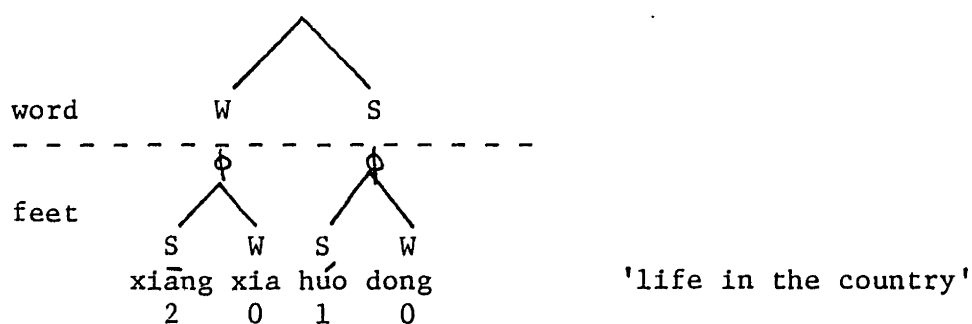
One of the little explored and particularly interesting questions concerns the interaction between tone and other suprasegmental phenomena such as stress. It seems likely that suprasegmental phenomena might be more closely related to each other than either is to the segmentals (at least when the suprasegmentals share certain phonetic characteristics, as do tone and stress). If such a relationship can be shown to exist it therefore offers confirmation of the suprasegmental analysis.

In chapter two this is discussed in detail, and I will only summarize my conclusions here. Working within the metrical theory as developed in Liberman and Prince (1977, henceforth L&P), stress is portrayed by means of binary trees whose nodes are labelled S (strong) and W (weak) such that two sister nodes always have different labels.

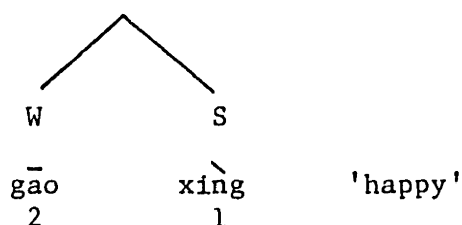
Main stress falls on the syllable dominated only by S nodes. Consider a Mandarin word with initial stress, and an unstressed neutral toned second syllable. It will have a stress tree that looks like this:



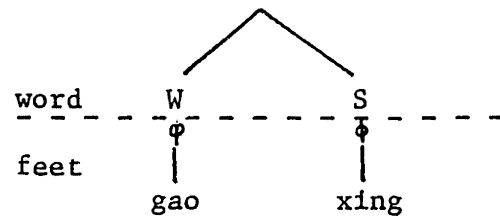
I argue that words like this are composed of single metrical feet, shown by ϕ , and such feet can in fact be combined into longer words:



When they are, the higher level of the tree -- the word level -- is WS rather than SW. This is exactly like the stress pattern in most bisyllabic words with full tone on both syllables, and main stress on the final syllable:



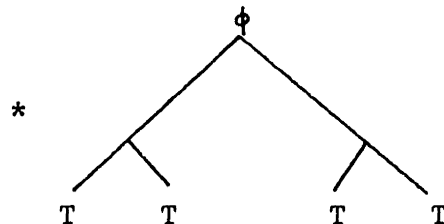
If this is also a word-level tree, presumably it is in fact built on feet just like the one for xiāngxiàhuódòng, but in this case each syllable is a foot of its own:



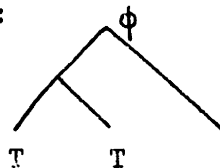
But we must now answer the question of what determines whether a syllable is a foot on its own, or whether several syllables are grouped into one foot. Whenever a foot includes two or more syllables, all but the first are neutral toned (i.e., toneless). This suggests that tone is what determines foot construction, and leads to the strong claim that feet are actually built on tones, rather than on syllables. Suppose that all fully toned syllables have two tonemes, such as HH, LH, LL, HL. These can be viewed as a branching structure:



Since, by universal conventions (see Halle and Vergnaud 1978) trees cannot branch in two directions within a single level, two fully-toned syllables cannot be joined into a single foot:



but one fully-toned syllable and one toneless (and therefore non-branching) syllable can be:



If we therefore assume that Mandarin metrical feet are left-branching and maximal, all toneless syllables will have to be attached leftwards into feet, but each fully-toned syllable will constitute a foot in itself.

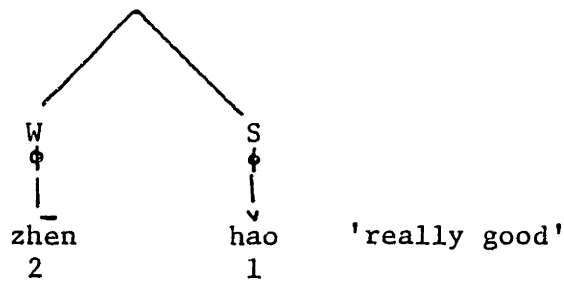
Floating tones (see 0.1 (b)) are taken into account by the tree construction rules too: since they are normally single tonemes they are treated as non-branching like toneless syllables, and attached leftwards in a foot. This leftwards metrical association also conditions their later leftwards tone-syllable association.

The foot as a unit plays several roles. Firstly, it is the domain of tone in the sense that there is a limit to the number of tones that are permitted in a single foot. Under certain circumstances (despite what has been said above) a foot can be constructed over two fully-toned syllables, but when this happens the tones of the second syllable are always eventually deleted (see also Safir (1979: 102), who shows that the foot is the domain of tone assignment in Capanahua). In verse the foot is also the domain of tone: two syllables in the same metrical foot will always be of the same tonal category.

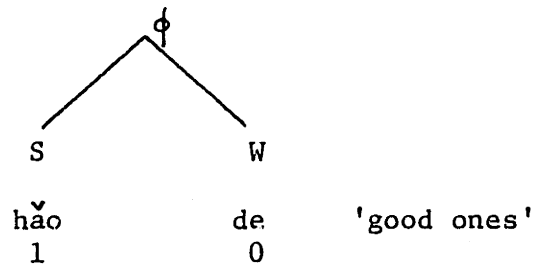
The second prominent role of the foot comes in the way it conditions tone sandhi rules. A number of sandhi rules group together two environments: pre-pausal position, and before a neutral tone. Given a level of the foot these two environments can be collapsed into "any position dominated by a terminal S node," for the following reason.

Pre-pausal position is always final position in the word (or actually phrase) tree. This tree is labelled WS, and assigns final

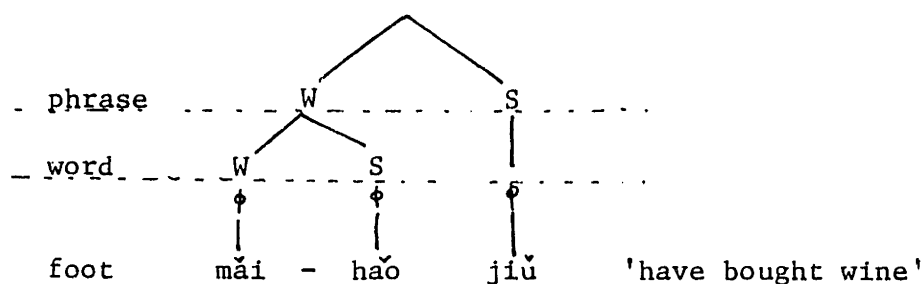
stress:



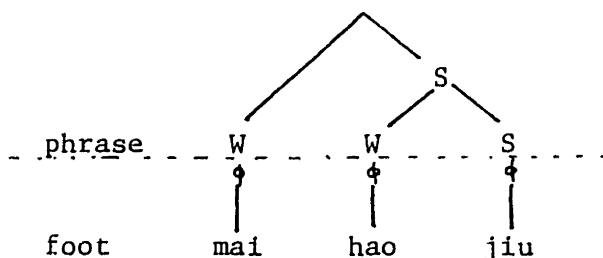
Pre-pausal position is thus always dominated by an S node. As for before a neutral tone, this will always be the first syllable of a foot, given the above discussion. Feet are labelled SW, and the first syllable is thus always S:



There is an interesting by-product of this argument: it appears that Mandarin has only two levels of metrical structure, feet and phrase levels. The word level is in fact only a special case of the phrase level, which immediately explains why word stress and phrase stress are identical in Mandarin (unlike English). The argument is based on the way the Mandarin third tone behaves. Before a pause or a neutral tone (i.e., when dominated by a terminal S node) it acquires a rise as a result of a rule which inserts a H tone. Elsewhere it is low level. Now if there were three levels of structure -- foot, word and phrase -- there would be phrases like the following:



Such a phrase has two terminal S nodes, one over each of the last two syllables. The sandhi rule ought therefore to insert two H tones, producing rises on both the last two syllables. In fact this does not happen: only the last gets a rise. This is exactly what we would expect if there is only a phrase level, and thus only one terminal S:



These two trees also differ in their predictions about secondary stress, and only the second correctly places it on the initial syllable.

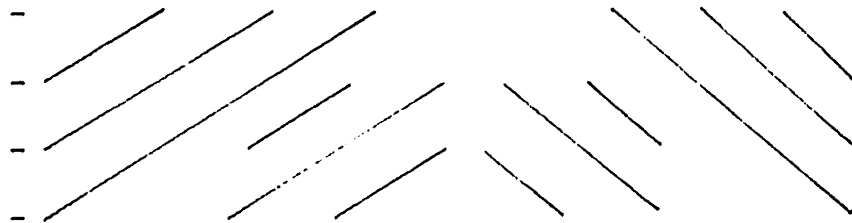
I conclude that word-level trees are not universally present, but only one available option. In languages that lack them word stress and phrase stress will always be identical, as in Mandarin.

The close interdependence of stress and tone is exactly what would be expected if both were suprasegmental, and provides further confirmation of the conclusions drawn from examining the purely tonal phenomena of stability and floating tones.

0.6 A Feature System For Tone

I argued earlier that contour tones must be represented by sequences of level tones, but this still leaves a number of questions unanswered. The level tones themselves will require a set of features, and the right set must predict the types of tonal inventory that are found in language (the evolution of tonal systems) and the natural classes of tones. These questions are discussed in detail in chapter three, and I summarize here.

It appears to be the case that the maximum underlying tonal inventory in any language is four level tones, and two tones of each contour (i.e., two rising, two falling, etc.). If contour tones are made up of level tones, once four levels are defined six tones of each contour can automatically be made up:



It is therefore hard to explain why no more than two of each contour are ever found. All previous feature systems run into this problem, and can only deal with it by placing an ad hoc condition on the system to block certain combinations.

The feature system I propose differs from earlier systems in that the two binary features which together define four levels are hierarchical: one of them, the Register feature, is 'dominant' and splits the pitch range of the voice into two halves. The second

feature, Tone (with an upper-case T), then further sub-divides each half to give a total of four levels:

<u>Register</u>	<u>Tone</u>
+ Upper	+ High (H)
	- High (L)
- Upper	+ High (H)
	- High (L)

It is claimed that each of these features is independently autosegmental (just as different vowel features may form separate tiers). However, only the feature Tone may occur in sequences underlyingly: Register remains constant over the syllable. The main effect of this is to restrict the basic inventory of tones to no more than two of any given contour:

+ Upper	L H	Rising
	H L	Falling
- Upper	L H	Rising
	H L	Falling

This system therefore satisfies the first requirement: it correctly categorizes the possible tonal inventories of languages. Note that it is quite possible that additional contour tones may arise in the course of a derivation as a result of derived sequences of Register over a single morpheme: the claim is only that the underlying contrast is limited to two tones of the same contour.

This feature system also satisfies the second requirement: it

allows a simple characterization of the historical development of tones. It is widely agreed that Chinese (and many other languages) underwent a process by which the voicing distinction on initial consonants was transformed into a tonal distinction, doubling the number of tones. Roughly speaking, voiced initials lowered all tones, and this is very simply dealt with in a Register system, since it is exactly what would happen if the feature Register were being brought into play for the first time: after voiced initials the tones become [- Upper], the rest (by default) [+ Upper]. It turns out, as I explain in detail in 3.1.1, that no other feature system can deal with these simple facts since they all lack a feature which has the necessary effect of raising/lowering all tones (while the remaining feature(s) express a simple two-way height contrast).

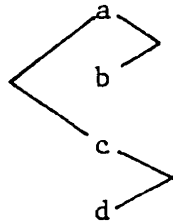
As far as the third criterion is concerned, correctly defining natural classes, it is easy to see that the register system makes very interesting predictions. Firstly, the independence of the two features makes a claim about the independence of general pitch level (given by Register) and melody (given by Tone). It is therefore possible to group together similar melodies, even when they differ in register. All rising tones have the Melody LH, and all falling tones the melody HL: each of these is thus a natural class. Notice that this is not so in most other feature systems that decompose contour tones: there is no sense in which the two rising tones MH and LM are a natural class. It is possible to see these kinds of groupings in action in the phonology. In Amoy (see section 5.3) there is a rule which dissimilates the first of two L Tones:

$$L \longrightarrow H / \text{---} L\#$$

This rule pays no attention to Register; the effect is therefore to change all tones with the melody LL to tones with the falling melody HL. In the Upper Register the change is 33 \longrightarrow 53, while in the Lower Register it is ll \longrightarrow 2l. It is noticeable that the vast majority of rules are of this kind: they act only on a single tier, irrespective of the specification of the other tiers.

The independence of the two tiers has more dramatic consequences: the general predictions of autosegmental phonology can now be made with respect to each tier. In particular, each tier should show stability during changes affecting another tier, and lexical entries should be able to lack one or more tiers. These predictions are explored in detail in chapter four, and shown to be correct. For example, Mandarin has various suffixes which always show up in the so-called 'neutral tone'. If these are analyzed as consisting of segmentals and [- Upper] Register, but no melody, their actual pitch is immediately accounted for. If the preceding morpheme ends with a H Tone, the suffix becomes [[- Upper], H] after the H spreads to satisfy the WFC. The actual pitch is then around the 3 or 4 level. If the preceding morpheme ends in a L Tone, the suffix becomes [- Upper, L], with a very low actual pitch of about level 1. The fourth tone is a falling HL, and therefore follows this second pattern. All the other three tones end in a H (in this position), and therefore take the first pattern. For a full explanation see section 4.1.1.

are four tones that alternate in three ways:



This can easily be stated in a Register system as follows:

$$\begin{aligned}
 a \sim c &= [+ \text{Upper}] \sim [- \text{Upper}] \\
 \left\{ \begin{array}{l} a \sim b \\ c \sim d \end{array} \right\} &= [+ \text{High}] \sim [- \text{High}]
 \end{aligned}$$

Other feature systems cannot normally account naturally for the $a \sim c$ alternation.

This phenomenon is not apparent in Chinese languages, which for the most part have only three level tones. Instead one observes an interesting ambivalence about the mid tone, which sometimes patterns with the high tone and sometimes with the low one. This is unsurprising, since it shares one feature with each no matter what its features, which may be either [+ Upper, L] or [- Upper, H]. The ambiguity of representation for the mid tone has some interesting consequences. Firstly, it suggests that one might find cases of phonetically identical mid tones that differ phonologically. Such a case is the Coastal dialect of Amoy, in which phonetically identical contextual tones have different forms before the suffix a⁵³ and are therefore presumably distinct phonologically (see section 5.3). Secondly, it predicts that if a three-tone language has a flip-flop

rule which switches the high and low tones, the mid tone will be phonetically unchanged. This is because such a flip-flop rule must be an alpha-rule affecting both Register and Tone:

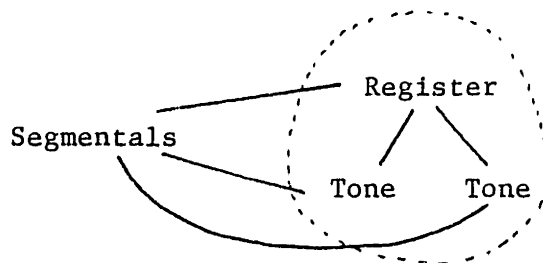
$$\begin{bmatrix} \alpha \text{ Upper} \\ \beta \text{ High} \end{bmatrix} \longrightarrow \begin{bmatrix} -\alpha \text{ Upper} \\ -\beta \text{ High} \end{bmatrix}$$

The effect of this rule on a mid tone will then be as follows:

$$\begin{bmatrix} + \text{ Upper} \\ - \text{ High} \end{bmatrix} \longleftrightarrow \begin{bmatrix} - \text{ Upper} \\ + \text{ High} \end{bmatrix}$$

(in either direction). In a three-tone system these two representations are equivalent phonetically, so the mid tone will be unchanged.

There is one last point to be made about the feature system proposed here. Although Register and Tone are independent tiers, they are clearly more closely associated to each other than either is to the segmentals. This may be partly a general property of suprasegmentals, especially those sharing some phonetic characteristics, but it is probably rather deeper. It seems likely that a more accurate diagram of the relationships between the various tiers would associate Register and Tone with each other, as well as with the segmentals:



The nature of the relationships within the encircled section has been

the concern of two recent papers, Clements (1979) and Huang (1979). Huang is concerned mainly with downdrift and downstep, whereas Clements tries to integrate the features of tone and the phenomenon of downdrift into a single system; both writers work within the metrical theory of Halle and Vergnaud (1978), and reach very similar conclusions (although there are differences of detail) on the hierarchical nature of downdrift.

It may also be worth noting in passing that there is some evidence that the Register distinction is acquired first and lost last, whereas the Tone distinction (i.e., the melody) is acquired later and lost more easily. Hashimoto (personal communication) tells me that both normal and auditorily handicapped children acquire level tones before contour tones, and T'sou (personal communication) tells me that Cantonese aphasics lose certain contour tones before the level tones.

0.7 Some Formal Properties of the System

In general terms it is clear that both Register and Tone exhibit behaviour of the kind predicted by Goldsmith's Well-Formedness Condition, but there are certain points where the Chinese facts suggest modifications and extensions of the theory.

First let us consider the form of the lexical entry and the language-specific initial tone-association rules. Logically there are two possibilities for the lexical entry: either the various tiers are entirely unassociated in the lexicon, or some association lines are already present. In the languages discussed here I will assume that the

first possibility is the right one -- the tiers are entirely unassociated in the lexicon. This allows for a very natural division of reduplication rules in Mandarin into those which apply before association and reduplicate only one tier, and those which apply afterwards and reduplicate the entire morpheme. It also accounts for the fact that the principles of association in each language are entirely regular: one does not find lexical entries which differ only in the placing of their association lines. However, neither of these arguments entirely rules out the possibility of association lines being present in the lexical entry: the regularity could be the result of a lexical redundancy rule, and reduplication could act on one tier alone even after association. In opting for the absence of such lexical association lines I am simply taking the view that this is the unmarked case for tone (as it certainly is in African languages) and that the burden of proof lies on anyone who wishes to show their presence.

One obvious way in which languages may differ is in the form of their initial tone association rules. These most commonly associate single tones with single tone-bearing units, and the most usual rule associates the first tone with the first tone-bearing unit, within a particular domain. In Shanghai the rule is of this type, and the domain is the syllable (or perhaps the morpheme -- the two are hard to distinguish in Chinese). The remaining association lines are inserted by the WFC. For example, consider a morpheme with high falling tone, 53. In isolation the initial tone-association rule of Shanghai will produce:

thy 'heaven'
 5 3

and the WFC will then associate the remaining tone:

thy
 5 3

When this morpheme begins a longer word, subsequent morphemes lose their own tones (by the regular word-formation processes of Shanghai).

Initial association produces:

thy sang 'naturally' sɿ sang cɿ 'a bastard'
 5 3 5 3

and the WFC results in:

thy sang sɿ sang cɿ
 5 3 5 3

The effect is to produce a falling tone on the isolated morpheme thy⁵³ but a level tone on this morpheme followed by a mid tone on subsequent morphemes when it begins a longer word. See section 5.2 for a full treatment.

Some languages have rather different initial association rules, however. Consider the equivalent facts in Mandarin. A morpheme with high falling tone, such as mai⁴¹ 'sell', retains this fall even when it is followed by toneless syllables. These syllables themselves end up low toned. The output must therefore be:

mai de 'the seller'
 ^
 4 1

This suggests that in Mandarin, unlike in Shanghai, all the tones of a morpheme are associated with the segmentals of that morpheme by the initial association rule. The WFC only comes into operation when additional, toneless syllables are added.

After initial association the WFC is the central core of the phonology. It turns out that Goldsmith's formulation requires certain modifications as suggested by Clements and Ford (1979). The most important of these concerns the requirement that every tone be associated with some tone-bearing unit. It turns out that this is much too strong, and that in fact tones may remain unassociated unless associated by a language specific rule. In Shanghai the tone of the second syllable is idiosyncratically left undeleted in certain circumstances. Since only the first tone of each morpheme is associated by the initial tone-association rules of Shanghai, the second tone of the first morpheme is therefore unassociated:

 \$ \$
 | |
 L H L

In such cases the unassociated tone is simply not realized, suggesting that Clements and Ford's reformulation is correct. Notice however that in most Chinese languages there are language specific rules which do require association of all tones, and the association is always leftwards in the languages considered here. This is true for

Cantonese floating tone, Amoy floating tone in reduplication, Mandarin inserted H tone, and word-final free tones in Shanghai.

The leftwards direction of association is not accidental, and it is related to another set of facts. Clements and Ford (p. 182) point out that there is a considerable amount of indeterminacy in the WFC, as mentioned in section 0.5. For example, in the configuration below the WFC can be satisfied by either (a), (b), or (c) (and both (a) and (b) involve minimal additions, and cannot be distinguished on grounds of simplicity):

$$\begin{array}{ccc} \$ & \$ & \$ \\ | & & | \\ T & & T \end{array}$$

(a) $\begin{array}{ccc} \$ & \$ & \$ \\ | & \cdot & | \\ T & & T \end{array}$

(b) $\begin{array}{ccc} \$ & \$ & \$ \\ | & \cdot & | \\ T & & T \end{array}$

(c) $\begin{array}{ccc} \$ & \$ & \$ \\ | & \cdot & | \\ T & & T \end{array}$

The amendment proposed by Clements and Ford (their convention 3, p. 186) is as follows:

(-*)

$$\begin{array}{ccc} \$ & & \$' \\ | & \cdot & \\ T & & Q \end{array}$$

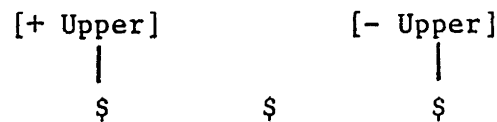
(mirror image)

(I have substituted the syllable \$ for their tone-bearing unit.) This expands in the unstarred (i.e., unaccented) cases that concern us here into two ordered cases:



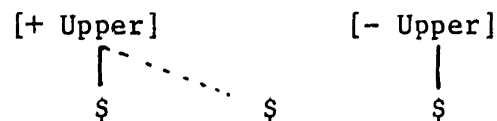
The claim is therefore that spreading from the left will be the dominant case, with spreading from the right happening only when there is no tone to the left.

This claim is borne out by data in Shanghai (see section 5.2). On tri-syllabic words the final syllable has lower Register associated with it. If the initial syllable has upper Register, the picture after initial association is as follows:



(ignoring the Tonal melody).

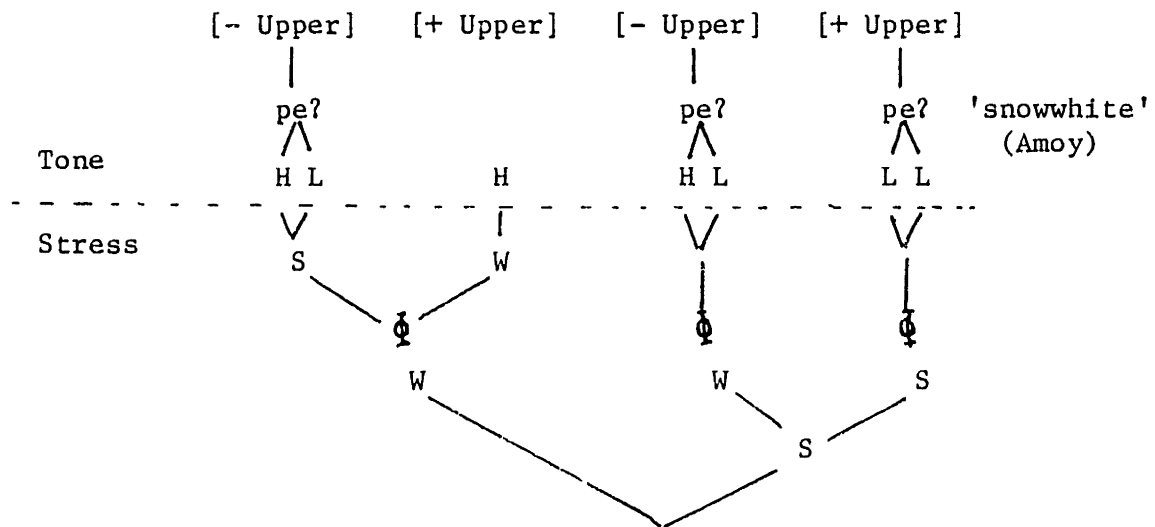
It turns out that the middle syllable surfaces with [+ Upper] Register, so clearly the spreading is from the left and not from the right.



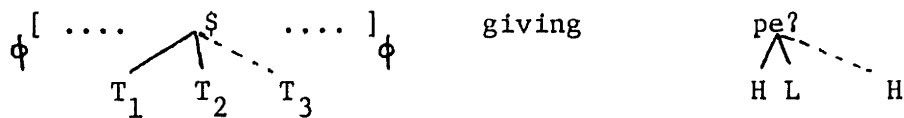
I would now like to go one stage further and suggest that Clements and Ford's convention may be derivable from other facts about the languages involved, and therefore need not be stated explicitly. The relevant facts are those of stress, and they also explain why floating tones are attached to the left and not to the right in these

languages. I would like to suggest that metrical structure is the conditioning factor for tone association, so that multiple associations are always foot-internal, and conform to the direction of branching of the foot.

I have claimed that stress trees are built over Tones. If this is taken literally it means that floating tones will be taken into account by the stress rules. Now floating tones consist of single Tones, not sequences. They are therefore unbranching nodes, and will be adjoined to the left in a foot by the normal foot-formation rules of the language:



It is then only necessary to assume that the rule of tone association applies foot internally:

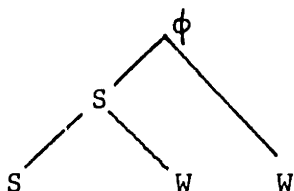


(and before the deletion of any tones dominated by W nodes).

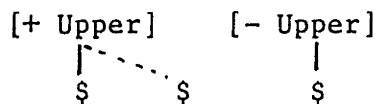
Something further must be said to account for Register associating

to the left, but it seems reasonable to suggest that no language would associate Tone one way and Register the other, presumably because they form a sub-unit of their own.

The stress facts also explain why tone has a preference for spreading in one direction. Returning to the Shanghai case, Shanghai words are single feet which are left branching, assigning initial stress.



If the 'tree' created by tone spreading is to be congruent with the metrical structure it must therefore involve branching on the left rather than on the right:



This approach to the asymmetry in spreading makes strong and therefore refutable predictions about the correlation between stress and tone spreading in languages. It remains to be seen whether these predictions are borne out.

The last question to be discussed concerns the possibility of contrasts between single tones and sequences of identical tones, such as L versus LL. Leben's Obligatory Contour Principle required alternation of tones underlyingly, thereby disallowing LL and HH sequences. Goldsmith, on the other hand, admitted an underlying

contrast, but claimed that the contrast was neutralized phonetically. He further proposed that the simplest grammar would always use a single tone rather than a sequence, L rather than LL.

I argue that for stress reasons the reverse is the case in Chinese: all morphemes have two Tones associated with them, different or identical. Mandarin, for example, has the four melodies LL, HH, LH, HL (all the possible two-Tone sequences). I see no reason to rule out the possibility of a contrast between, say HH and H underlyingly (and in fact floating tones in Mandarin and Cantonese are analyzed as single H Tones), but it does not play a crucial role in any language discussed here. However, such contrasts do arise in the course of the derivation (for example, in Amoy, see section 5.3) and may then differ in whether or not they undergo subsequent rules. At the phonetic level I would claim, with Goldsmith, that the level tones H and HH, L and LL, are not distinct. This is certainly true in Amoy. However I would depart from Goldsmith and 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. The transition from L to H pitch occurs at a different point in the syllable in such cases (see section 4.1.2 on Mandarin intonation), whereas there is no transition point to signal the difference between plain HH and H.

This completes the summary of the major findings of this thesis: the interested reader will find detailed data and arguments in the body of the text.

1. ARGUMENTS FOR AN AUTOSEGMENTAL ANALYSIS OF CHINESE LANGUAGES

In this chapter we will give examples of the sort of tonal behaviour predicted by autosegmental phonology and actually found in various Chinese languages. The predictions fall into two groups (see also section 0.1.2). The first set involve the independence of the two levels as lexical entries, and therefore the existence of morphemes lacking one or other component. These are exemplified in 1.1 and 1.2. The second set involve stability: the operation of phonological rules on one level only (usually deletion rules), and also the spreading of a melody to all available syllables. These are discussed in sections 1.3-1.5. In section 1.6 we discuss a putative case of a language with segmental tone, and conclude that it is not in fact possible to tell whether tone is autosegmental or not in that language (Zahao) and that therefore it is still possible to maintain an autosegmental analysis as the unmarked case for tone. And we conclude this chapter with a discussion of Clark's (1978) proposal for dynamic tonal markers interspersed among the segmental layer, and show that her analysis fails on both principled and language-specific grounds.

1.1 Segmentless Morphemes

1.1.1 Cantonese Changed Tone

The data on which the analyses of Cantonese in this thesis are based are taken largely from Hashimoto (1972), supplemented by data from Whitaker (1955-6) and Chao (1947). Other useful sources include HYFYCH, HYFYGY, Kao (1971) (the last is a phonetic study).

Cantonese is a Yue dialect, and the name Cantonese is confusingly used for both the dialect spoken in Canton and Hong Kong, and as a synonym for Yue to denote the language group as a whole. As used here it refers to the specific dialect. The Yue group has retained the full range of historically present final obstruents -p, -t, -k, and in Cantonese these stopped syllables can bear only three tones underlyingly, all level and shorter than the unstopped variants. Assuming that the length is predictable from the presence of the final stop, Cantonese may then be considered to have a total of six (or seven) underlying tones. The uncertainty comes because one sandhi alternant (/53/ → 55) is apparently in the process of becoming a separate underlying tone, since a few morphemes no longer exhibit the alternation but occur always in what is usually considered the derived tone, /55/.

Under certain morphological conditions all tones have a set of special alternants called changed or modified tones (bìan yīn). For two of the tones this alternant is a high level tone; for all other tones it is high rising. Below we give the underlying tones, followed by their alternative forms:

$$\begin{array}{r}
 55 \text{ (5)} \\
 53 \quad \left. \vphantom{\begin{array}{l} 55 \\ 53 \end{array}} \right\} \longrightarrow 55
 \end{array}
 \qquad
 \begin{array}{r}
 44 \text{ (4)} \\
 33 \text{ (3)} \\
 22/21 \\
 35 \\
 24 \quad \left. \vphantom{\begin{array}{l} 44 \\ 33 \\ 22/21 \\ 35 \\ 24 \end{array}} \right\} \longrightarrow 35
 \end{array}$$

The single digits in parentheses are the traditional way of denoting the short stopped tones, which we take to be the conditioned

variants. Notice that even the stopped syllables can acquire the rising contour tone, even though one of them consists of a single mora.

There is some disagreement about the phonetics of the derived rising tone, which is normally written [35] or [*35]. Hashimoto states that it is a 'high rising tone, similar to the Yang Shang tone (i.e., underlying /35/)' (p. 93). Chao considers it to be slightly distinct, starting somewhat lower and being rather longer, say [25:].

The most interesting remarks are those made by O'Melia (1939, p. xxiii):

"The real tone of the first word is given, though barely touched, after which the voice sails off like a kite, often sounding exactly as though it were an upper rising tone, and sometimes rising higher than the upper rising."

Dyer Ball, quoted in Whitaker (1956: 190):

"...the distinction between these five or more rising variant tones has not been pointed out or clearly defined and they have all been considered by many as one and the same tone."

and Whitaker (p. 193):

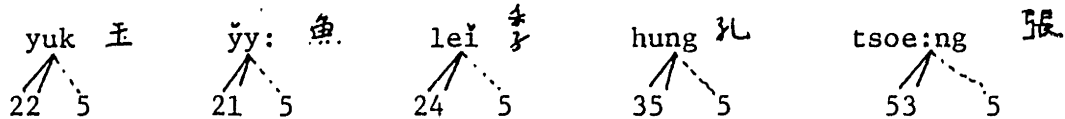
"Each of the...modified tones begins from the starting pitch of the tone it modifies and takes it up to 5."

T. Cheng (1973) has shown that the same is true for the closely related dialect of Taishan.

Now if we examine again which of the tones becomes high level [55], and which becomes high rising [35] we can make the following statement. If a tone begins at a high level (i.e., level 5), then it becomes a high level tone. If it begins at any lower level, then

it becomes a rising tone ending at level 5.

All this can be explained if what is happening is that a high level floating tone is suffixed in certain morphological environments. To duck the feature issue for the time being we will use digits. Consider five cases:



The WFC will require association of the floating tone, and we will then derive tones that, in the first four cases, start at their normal starting point and rise to a high level; the presence of a final stop is irrelevant. In the fourth case we derive a tone which starts and ends high, but dips in the middle. This dip is apparently not realized (although it is in Taishan (T. Cheng 1973)), which should not surprise us since there is undoubtedly a limit to the amount of phonetic detail that can be realized on one syllable.

The simple assumption of a floating high tone -- a natural possibility in autosegmental theory, but a real problem in a segmental approach -- then accounts for exactly the subtle differences in the forms of the modified tone observed by O'Melia, Dyer Ball, and Whitaker, including why the 35 tone alone does not change, and why some tones become high level. It is to be hoped that phonetic research will confirm these distinctions.

What kind of morpheme is this floating tone? It has a number of uses. The most common is a kind of diminutive or familiar usage, so that many household objects are normally referred to in this tone:

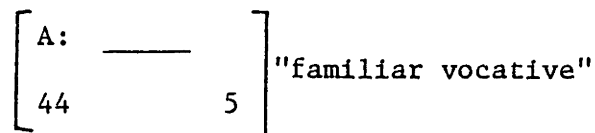
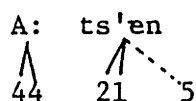
$\check{y}y:^{21} \rightarrow \check{y}y:^{*35}$ 'fish' 魚 but $yy:^{21}$ $hA:^{53}$ $ha:i^{24}$ 'fish, shrimp, crab' 魚 蝦 蟹
 $ts'A:k^3 \rightarrow ts'A:k^{*35}$ 'thief' 賊 but tou^{33} $ts'A:k^3$ 'thieves and robbers' 盜賊

Which morphemes acquire the floating tone, and in which combinations, is entirely unpredictable, and must be marked in the lexicon. Note that since this is so one could in theory enter each morpheme with two tones, different ones for each context. But this would make it complete chance that we find no words which occur with, for example, [21] in some combinations and [24] in others. It would also fail to explain why no morpheme alternates [5] and [35], for example.¹

In any case there is a productive use of the modified tone that is not susceptible to simple listing in the lexicon, since the environment is predictable. There are two morphemes commonly prefixed to family names in familiar use. After these prefixes the names assume changed tone:

$A:^{44}$	$ts'en^{21}$	\longrightarrow	$A: ts'en^{*35}$	阿陳
lou^{24}	lei^{24}	\longrightarrow	$lou^{24} lei^{*35}$	老李
$A:^{44}$	$tsoe:ng^{53}$	\longrightarrow	$A:^{44} tsoe:ng^{*55}$	阿張
$A:^{44}$	$ii:p^{22}$	\longrightarrow	$A:^{44} ii:p^{*35}$	阿葉

The correspondence between underlying tone and changed tone forms is identical to that in the earlier examples, but these names all occur frequently with their basic tone. The assumption of a floating high tone immediately explains the observed forms, and we see that the familiar affixes are actually made up of a prefix and a suffix, the suffix being purely a high tone.



The changed tones can be explained very simply in an autosegmental framework. In a segmental approach things are not so straightforward. First it is not clear how floating tones could be described at all. Secondly, when the output is a contour tone on a single mora it would be necessary in a level-tone analysis to claim that in fact there were two moras phonetically and write a rule inserting a mora in just those cases. There is some slight evidence that such syllables are lengthened a little in changed tone, but not, as far as I know, to the point where they are as long as underlying bi-moraic syllables.

If on the other hand one adopts a contour tone analysis like Wang's it is necessary to have two separate rules:

$$\left[\begin{array}{c} +\text{high} \\ -\text{rise} \end{array} \right] \longrightarrow [-\text{contour}]$$

$$[-\text{high}] \longrightarrow \left[\begin{array}{c} +\text{high} \\ +\text{rise} \end{array} \right]$$

These two changes are entirely unlinked, and the fact that [+high, +rise] alone is unchanged is unexplained. The context would have to be morphological rather than phonological, and in order to state the second rule in the fairly simple form given here it is necessary to assume that even the 44 tone is [-high].

We conclude, then, that only an autosegmental analysis provides an elegant explanation for the facts of Cantonese changed tone.

1.1.2 Mandarin Reduplication

The reduplication facts are included in this section because, as we shall see, not only is the underlying tone not reduplicated, but a floating high tone is added (sometimes accompanied by retroflexion). As argued above, in a segmental analysis there is no simple way of explaining the occurrence of floating tones of this kind. The arguments that follow would also belong by rights to section 1.2, since the result of the morpheme-level reduplication is a toneless morpheme, a possibility predicted by autosegmental phonology.

In a segmental analysis of tone nothing should be able to separate the tone from the rest of the morpheme. It follows therefore that reduplication of a morpheme must reduplicate both segments and tone. In an autosegmental phonology the two levels may be unassociated in the lexicon. If therefore a morphological rule operates on that lexical entry it might in theory be able to reduplicate only one of the two levels. Depending on one's theory of morphology, rules might be expected to differ in what they take as input, and to what extent the internal structure of the lexical entry is available for inspection. I would like to propose that Mandarin has two different types of reduplication, one of which applies at the word level, and cannot survey the composition of that word, but simply reduplicates it as a unit, tone and all. The second kind of reduplication applies at the morpheme level, and therefore has the lexical entry as input. This kind of reduplication chooses only the segmentals and reduplicates those but not the tone. This second kind of reduplication is only conceivable in an autosegmental phonology.

Notice, interestingly, that one result of this assumption will turn out to be that a boundary distinction proposed for Mandarin loses its only motivation.

The standard analysis of Mandarin reduplication is due to C-C Cheng (1973). Without the apparatus of autosegmental theory he assumes that all reduplication is of the first type that reduplicates tone as well as segmentals. Certain differences in tone sandhi therefore force him to postulate an otherwise unmotivated distinction between + and #.

First let me summarize the facts of the more productive word-level type of reduplication. Two categories are involved. Classifiers may be reduplicated with a distributive meaning (Chao 1968: 202):

- | | | | | |
|-----|-----|-------|-------|---------------|
| (1) | 張 張 | zhāng | zhāng | 'every sheet' |
| | 寸 寸 | cùn | cùn | 'every inch' |
| | 种 种 | zhǒng | zhǒng | 'every kind' |

Notice that in the first two examples the two occurrences are in every way identical, suggesting that the rule copies the entire morpheme. In the third case there is an apparent difference in tone. However, this is the same regular tone sandhi that occurs on any sequence of two third tones to change the first one into a second tone:

- | | | | | | |
|-----|-----|------------------|----|--------------------|-----|
| (2) | 这种狗 | zhèi zhǒng gǒu | or | liǎng zhǒng shū | 两种书 |
| | | this kind of dog | | two kinds of books | |

So again the assumption that the entire morpheme is copied will produce the observed output. In some cases, or for some speakers

(e.g., Cheng 1973:72), the second occurrence shows up in neutral tone. It has been argued that neutral toned syllables are unstressed syllables with no tone of their own, so superficially examples such as:

- (3) 个个 gè gè 'every one'
 种种 zhǒng zhǒng 'every kind'

might suggest, in contrast to the earlier examples, that the copying rule copies only the segmentals, and not tone. However, observe the tone on the first syllable of zhǒng zhǒng: it has apparently undergone the tone sandhi rule which changes a third tone to a second tone before another third tone. We must thus assume that the third tone was also copied, and later deleted. All these facts follow if we assume a rule reduplicating the whole morpheme, plus an optional tone deletion rule, syntactically conditioned (the same rule that applied to object pronouns, for example)(see section 2.1.1).

Now consider the second type of productive reduplication. This is a process that reduplicates verbs, and carries a tentative meaning (Chao 1968:204):

- (4) 看看 kàn kàn 'just look'
 想想 xiǎng xiǎng 'just think' (from third tone xiǎng)

The third tone sandhi applies in the second example given here too, so we can conclude that this process is exactly parallel to the distributive reduplication discussed above.

We will now turn to the more interesting cases. There is a fairly productive process which reduplicates adjectives to give a

more 'vivid' meaning (Chao:205). When mono-syllabic adjectives undergo such reduplication the second syllable takes on the first tone (high level), and in more informal speech the retroflex suffix /r/ is also added in many dialects. Literary words would normally lack this suffix, and also the tonal change.

- (5) cháng chāngr de 'often' 常常儿地
 jiàn jiàn de 'gradually' (lit.) or 渐渐地
 jiàn jiānr de (informal)

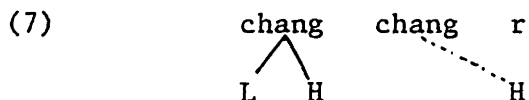
These reduplicated forms are finally stressed, and we can see from the forms without tone change that there is no deletion of tone in these cases. Now if we assume that the entire morpheme is reduplicated as in the above examples, a form such as cháng will reduplicate to:

- (6)
- | | | |
|----------------|----------------|---|
| chang | chang | r |
| / \
L H | / \
L H | H |

(I have implicitly suggested that there is a morpheme consisting of retroflexion and high tone. In fact, in many dialects the suffix is H tone only, with retroflexion absent or optional. Elsewhere in the grammar we find retroflexion with no tone change, and we also find what is historically probably the same morpheme 儿 ér with a high rising second tone, meaning 'son'. Clearly they are no longer a single morpheme, so attempting to unite them by positing a high rising tone for the suffix r seems an unnecessary complication).

In order to account for the surface form with plain high tone on the second syllable we would have to introduce ad hoc deletion rules deleting all the tones attached to the second syllable. If on the

other hand we assume that reduplication does not copy the tone in the first place, we have:



and the WFC will require association as shown by the dotted line.

Now consider the reduplication of third tone adjectives:

(8) 好好地 hǎo haor de 'good, well'

Notice that no third tone sandhi has applied. If the reduplication rule did not copy the tone, then there is no more to be said. However, if the tone had also been copied, rule-ordering would have to be invoked to account for the lack of sandhi: in other words the deletion rules would have to precede, and therefore bleed, the sandhi.

It should be clear, then, that the assumption that this type of reduplication does not copy the tonal component simplifies the phonological component. We would expect such lexical reduplication (lexical since it obviously preceded association of tones and segments) would show more idiosyncracies than does purely syntactic or word-level reduplication, and this is indeed the case. In contrast to syntactic reduplication, in which all classifiers² and (as far as I know) all verbs for which the tentative aspect is semantically acceptable can be reduplicated, reduplication of adjectives is subject to certain restrictions, at least for bisyllabic forms. Monosyllables reduplicate quite freely, although some, such as huai "bad", do not reduplicate, although the existence of adjectives which reduplicate with their tone (like jian above) will require special marking. Such marking is even

more necessary to deal with bisyllabic adjectives. For such adjectives there are three possible reduplicated forms, but for any given adjective it is usually the case that only one is possible (or at most two). The patterns involve reduplication of either element, or of both elements:

- | | | | |
|-----|------|-------------------|----------------------|
| (9) | 崩崩脆 | bēng bēng cuì | 'crackling crisp' |
| | 冷冰冰 | lěng bīng bīng | 'icy cold' |
| | 快快乐活 | kuài kuài huó huó | 'happy and cheerful' |
| | 清清楚楚 | qīng qīng chǔ chǔ | 'perfectly clear' |

Most forms involving reduplication of the second element only, as well as many of the XXYY type forms, have no unreduplicated equivalent and must be entered in the lexicon in this form. What is more, each ordinary bisyllabic adjective will need an entry specifying in which of the possible ways it reduplicates. There is one last crucial difference between these reduplicates and verbal reduplicates. Notice that the adjectival forms here split up the word into its two morphemes and reduplication then acts on those. What it does not do is reduplicate the entire word:

- (10) * kuàihuó kuàihuó 'happy and cheerful'

This remains true even when the adjective has a meaning which is non-compositional, and therefore the derivation from component morphemes may not be clear to the speaker:

- (11) 規規矩矩 $\bar{g}u\bar{i} \bar{g}u\bar{i} j\check{u} j\check{u}$ 'well behaved and well mannered'
 規 $\bar{g}u\bar{i}$ 'compasses'
 矩 $j\check{u}$ 'carpenter's square'

Now contrast this to verbal reduplication, which is unable to break the word down into component morphemes:

- (12) 討論討論 $t\check{a}o\ l\grave{u}n\ t\check{a}o\ l\grave{u}n$ 'discuss a little'
 琢磨琢磨 $z\acute{u}o\ m\acute{o}\ z\acute{u}o\ m\acute{o}$ 'thinks over a little'

We never find:

- (13) * $t\acute{a}o\ t\check{a}o\ l\grave{u}n\ l\grave{u}n$ 'discuss a little'

If this is a syntactic word-level process this is of course what we would expect. For completeness I should mention one idiosyncrasy of these polysyllabic reduplicates. If the process were simply an extension of monosyllabic reduplication we might expect the high tone (and retroflexion) to show up on the reduplicated element. This is only partially true. The expected forms would be:

- (14) XXY * $b\bar{e}ng\ b\bar{e}ngr\ c\grave{u}i$
 XYX $l\bar{e}ng\ b\bar{i}ng\ b\bar{i}ngr$
 XXYY * $ku\grave{a}i\ ku\grave{a}ir\ hu\acute{o}\ hu\acute{o}r$

There is apparently a constraint on the insertion of retroflexion internally in these forms. This accounts for the observed forms in the first and last cases (see (9)) including the fact that cui does not change tone, but huo does. The lack of tone change on the reduplicated first syllable is probably deceptive. In fact these

surface with neutral tone, and we may thus assume that the tone did in fact change. The main difference occurs in forms whose last element has been reduplicated -- XYY or XXYY -- but does not change tone. One example was given earlier, qīng qīng chū chǔ and this word, like many others, in fact allows optional tone change on the final syllable: qīng qīng chǔ chū. These are of course simply the counterpart of the monosyllabic adjectives which fail to change tone, like jiàn, and can be handled in the same way, although they are more common.³ Notice that, at least when the base adjective is trochaic (i.e., "neutral toned" on the last syllable meaning unstressed and toneless), the occurrence of the basic rather than the changed first tone on the last syllable is considered literary (Chao:206) just as is the case with monosyllables like jiàn.

I should now like to turn to the second type of lexical reduplication. This class of cases was the one which caused Cheng to introduce the + boundary, and is therefore of particular interest. There is a fixed class of reduplicated nouns consisting mainly of kinship terms:

- | | | | |
|------|----|---------|-----------------|
| (15) | 弟弟 | dì dì | 'younger' |
| | 哥哥 | gē gē | 'older brother' |
| | 姐姐 | jiě jiě | 'older sister' |

These all have trochaic stress, so the tone of the second element does not surface. Presumably the speaker knows what it is, however, since the derivation of these items is transparent. Now notice that when a third tone morpheme is reduplicated, as in jiě jiě, no tone sandhi takes place. This is in direct contrast with non-reduplicated

lexical items that are otherwise identical:

(16) 小姐 xiǎo jiě → xiǎo jiě 'Miss'

Since the sandhi process normally occurs across words in a string, Cheng suggested that the class 'xiǎo jiě' had internal #, allowing sandhi to apply, but reduplicated words had only +, and sandhi therefore failed.

By now the reader will have seen what is coming: if lexical reduplication does not copy tone, there is no way that sandhi could apply in these cases and nothing further need be said.

Although this class of nouns is normally closed, mothers using 'baby talk' can extend it quite productively, and when they do so the facts remain the same:

(17) 狗狗 gǒu gǒu 'doggy'

Indeed, we can find at least one minimal pair. There is a verb that belongs to this closed class in some dialects:

(18) 痒痒 yǎng yǎng 'it itches'

In the same dialects there is a monosyllabic but homophonous verb meaning 'to cultivate'. If this undergoes the syntactic reduplication associated with the tentative aspect we find:

(19) 养养 yǎng yǎng 'cultivate a little'

That is, sandhi has applied in this case but not in the other.

To summarize, it appears that Mandarin has two different processes of reduplication. One applies at the word level, and

reduplicates freely without access to lexical exception markings, or access to sub-parts of the word: for this reason bisyllables are reduplicated as units, XYXY, and not as morphemes (XXYY). This type of reduplication copies the entire word, including its associated tone. The second type of reduplication applies in the lexicon; it is subject to exceptions and idiosyncratic specifications on the precise type of reduplication. Since it has access to the lexical entry it is able to analyze the word (sometimes etymologically-speaking wrongly) into morphemes and reduplicates these rather than the word. In the typical cases it does not reduplicate the melody, but only the segmentals: this is possible in a theory which considers tone and segmentals to be unassociated in the lexical entry, but associated later. We may now apparently dispense with any +/# distinction for Mandarin, since as far as I know the failure of tone sandhi in reduplicated nouns was the only motivation for such a distinction.

There is a further prediction of this analysis: reduplicated adjectives and nouns are considered lexical items, and should therefore be immune to any syntactic rules which might break them up. Reduplicated verbs and measure words, on the other hand, should not be so protected, and indeed certain things can appear between the two parts: the numeral yí 'one' for example, and the perfective marker le:

yáng yí yang 'cultivate a little'
 yáng le yang 'having cultivated a little'

1.1.3 Amoy Triple Reduplication

Amoy is a Southern Min dialect, a variety of which is spoken in Taiwan. The data on which the following is based is taken from a book

and two papers by R. L. Cheng (1977, 1968, 1973), who is a native speaker of the Tainan dialect of Southern Taiwan. A full analysis is included in chapter 5, but here we are interested only in one set of tonal changes.

In Taiwanese it is possible to reduplicate adjectives three times. This so-called triple reduplication is accompanied by a special set of tonal changes in the first syllable. Now Taiwanese tones, like those of many Chinese languages, always possess two allotones, one of which shows up in pre-pausal position (and before toneless syllables) and the other of which shows up before any other tone. For reasons that will become clear in chapter 5 it is necessary to assume that the pre-pausal (or isolation) form is basic and the contextual form derived. This general process presumably precedes the more specific processes that take place in triple reduplication, and we will take the derived contextual forms to be the input to the special rules. Below we give the underlying, contextual and special triple reduplicated forms of the tones:

Underlying	53	21	21?	54?	33	13	55
Contextual	55	53	54?	21?	21	33	
Triply Reduplicated	55	53	54?	35			

(Two of the tones end in a glottal stop; just as in Cantonese they can be taken to be conditioned variants of the similar longer tones, as will be discussed in detail in chapter 5.)

Let us give some examples of triply reduplicated forms.

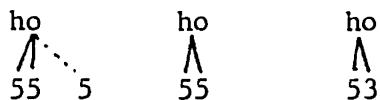
$pe?^{54} pe?^{54} pe?^{54} \rightarrow pe?^{21} pe?^{21} pe?^{54} \rightarrow pe?^{35} pe?^{21} pe?^{54}$ 白
 'snowwhite'
 $ho^{53} ho^{53} ho^{53} \rightarrow ho^{55} ho^{55} ho^{53}$ 'very good' 43

The tones on the left-hand side of the table undergo no special changes in the triply reduplicated forms, but the tones on the right-hand side all become rising to the highest tone level. This looks suspiciously like the situation with Cantonese changed tone, and indeed if we again postulate a floating high tone in these examples we will account for most of the facts. Using digits still, we have, after the contextual rules have applied:

$$\begin{array}{ccc}
 pe? & pe? & pe? \\
 \wedge & \wedge & \wedge \\
 21 & 21 & 54
 \end{array}$$

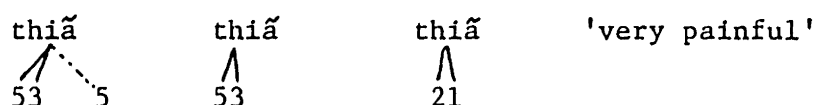
This will produce a rising tone; whether or not there is any distinction between this and the rise derived from syllables with a contextual 33 tone is not clear. No such distinction is mentioned in the relatively limited literature I have been able to find on this dialect. If there were such a distinction it is likely that it would be so small that its absence in the literature would not be surprising.

This assumption will also rightly account for the lack of any surface change in two of the tones on the left-hand side: [55] and [54]. Both of these are probably phonologically high level, and the final fall is a late phonetic side effect of the final glottal stop (see Zee and Maddieson (1979) for evidence of such an effect in Shanghai, and Fok (1974) for similar evidence in Cantonese). Suppose the floating tone is added here also. Then we have:



and of course no surface change will be observed.

The one case that is not correctly predicted is the failure of the high falling tone to become high level, which is what it did in Cantonese. Consider a word like the following:



We would expect the complex [535] to be simplified to its beginning and end points like in Cantonese, instead of which the final high 5 is not realized. It seems then that there is a language specific rule which simplifies this sequence to [53] for reasons that are not clear.

Apart from the last detail, a floating tone analysis of Amoy triple reduplication simply explains the observed alternants, including the failure of the high level tones to change at all. Notice that this is just such a situation where the WFC does not provide a unique statement of how to achieve well-formedness. Unlike in Cantonese, where the floating tone was (usually) word-final, or Mandarin, where it always was, here it is word medial. The fact is that it always attaches leftwards, and this must either be stated language specifically, or by general convention. Clements and Ford do not deal with this problem, since they take the position that free tones are not subject to the WFC unless so stated for each language, since the direction of association cannot, in their opinion, be predicted. We

will return to this in chapter 5.

In this section we have seen three cases of floating tones. In each case the data is simply accounted for once a floating tone is assumed. Interestingly, it is always a high tone, but this is probably an artifact of the fact that these three cases may share a common origin in the morpheme that preceded the modern Mandarin ér suffix. Floating tones are to be expected in autosegmental theory, but constitute something of a problem for a segmental approach. Furthermore, no alternative analysis suggests itself in any of these cases. We conclude that these cases provide a strong argument for autosegmental tone in these three languages.

1.2 Toneless Morphemes

1.2.1 Mandarin Suffixes

Most morphemes in Mandarin carry one of four lexical tones, repeated here for ease of reference:

媽	mā	'mother'	First tone. 55 high level
麼	má	'hemp'	Second tone. 35 high rising
馬	mǎ	'horse'	Third tone. 214 low, with rise pre-pausally 21 low level elsewhere
罵	mà	'to scold'	Fourth tone. 41 falling

However, a small number of morphemes never show up with any of these tones, but are always in what is called 'neutral tone'. Phonetically this means that the pitch of such morphemes is predictable from the tone of the preceding morpheme in the following way.

	Preceding morpheme	Neutral toned morpheme
1st	55	3
2nd	35	3
3rd	21	4
4th	41	1

All such morphemes are clitic-like suffixes or particles which never occur in initial position in a phrase. For example, the subordinating particle de can be attached to anything to make it a modifier, or an agent in the case of verbs.⁴ It could be attached to any of the above examples in theory to give:

māde	'mother's'	máde	'hemp's'
mǎde	'horse's'	màde	'the one who is scolding'

Other examples include the aspectual marker le 'perfective', which is a suffix, or the sentence particles like ma, which turns the preceding sentence into a question:

Tā niànle sānnián Zhōngwén. 'He studied three years of Chinese'
 He study-PERF three-years Chinese language

Tā niànle sānnián Zhōngwén ma? 'Did he study three years of
 Chinese?'

There are two possibilities for analyzing these morphemes. Either they carry a fifth tone of some kind that undergoes rather complex tone sandhi, or they carry no tone at all of their own, and their phonetic pitch is acquired by spreading from the preceding morpheme. The fifth tone hypothesis can be discarded immediately as unsatisfactory because there are morphemes that show up in one combination with one of the

four lexical tones, and in some combinations with neutral tone. For example:

意 思	yìsi	'meaning'
思 想	sìxiǎng	'thinking, thought'

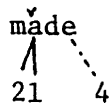
If the neutral tone is a fifth tone, then there is no reason why we should find morphemes which alternate between fifth tone and each of the other tones, but no alternations between, for example, first and fourth tone. If however it is just a toneless morpheme, then we need only say that tones are lost under certain circumstances.

We are therefore left with the assumption that these morphemes have no tone of their own. In that case let us examine what the theory predicts as the phonetic output. Consider first tones three and four. Tone three in citation form is [214]; when a neutral toned morpheme follows, the first syllable has tone [21], and the final syllable tone [4]:

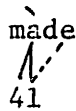


Since the WFC will require that the final tone, tone level 4, spread onto the toneless syllable it remains only to explain why it is not also associated with the first syllable in the final output. The reader can see, I think, that if de is toneless as proposed the pitch of de itself is immediately accounted for in these cases; in order to also account for the contour [21] rather than [214] on the first syllable we will have to request patience until the complete analysis of Mandarin is laid out in chapter five. To give a foretaste of the

solution, it is related to the fact that the third tone shows up as [21] before any other tone; this suggests that the 4 may in fact be inserted both pre-pausally, and in the neutral-toned environment. If it is inserted then it will of course associate one-to-one with the similarly unassociated de, and we will derive, correctly:



Now consider the fourth tone, which had the falling contour [41]. A following neutral tone shows up as a very low level tone, [1]. This follows directly from the well-formedness condition:



The other two tones, tones one, [55], and two, [35], trigger a mid-level tone on a following neutral toned syllable, although from what has been said so far one would have expected a high level tone:



It turns out, as we will see in chapter five, that this is because neutral toned syllables are toneless but not registerless, terms which will be explained in chapter three. They do indeed acquire high tone from the preceding syllable, but it interacts with their own lower register to produce a phonetically mid-level tone.

Although some details of the analysis must be postponed to a later chapter, it is clear that Mandarin possesses a group of morphemes that

are best characterized as toneless. In a segmentally based phonology this makes little sense; in a tone language all vowels must be tonally specified just as all vowels must be specified for backness. In an autosegmental phonology, however, it is to be expected, just as segmentless morphemes are to be expected.

1.2.2 Southern Min Neutral Tone

Taiwanese (R. L. Cheng 1966, 1968, 1973) and Chaoyang, another Southern Min dialect closely related to the better known Chaozhou, (Zhang 1979) are examples of two S.Min dialects which exhibit closely similar phenomena called, like the Mandarin example, neutral tone (qīng shēng), but that differ from Mandarin in a number of ways. Despite the differences it can be shown that neutral toned morphemes in both languages should be analyzed as toneless, just as they were for Mandarin.

As we said in section 1.1.3, S.Min dialects typically have two alternants for every tone, one of which occurs pre-pausally and the other elsewhere. This was in fact an imprecise statement: the pre-pausal variants occur in two positions:

- (i) pre-pausally
- (ii) before neutral tone, phonetically [ʔ]

There are two ways of unifying these apparently disparate environments, and either requires the assumption that neutral toned syllables are toneless. The first possibility is that (i) and (ii) are both situations where no tone immediately follows (within the domain, the definition of which is a separate problem). For example, the surname Tan¹³ still shows up in that citation form when followed by the neutral

toned suffix for 'family':

Tan¹³ ka 'The Tans, the Tan family' 陳家

Yet before any full-toned morpheme the [13] tone normally changes to a mid-level 33:

bə¹³ tsi¹³ → bə³³ tsi¹³ 不 錢

The second possibility is that the change is stress-conditioned (see section 2.1.2): Taiwanese and most Chinese languages have stress on the last non-neutral-toned syllable. The result is that syllables in isolation tones are those that carry main stress, while those in sandhi tones have lesser degrees of stress. However, this does not solve the question of why neutral toned syllables should be ignored for stress purposes. If they carry some special kind of tone, why should they alone be passed over by the stress rules? They could of course be marked with a special diacritic [-stress], and this was the tack taken by C-C Cheng (1973) in solving a similar problem in Mandarin. However, given an autosegmental analysis there is no need to have recourse to the power of diacritics for a solution. It is simply necessary to assume that these syllables have no tone, and that this itself conditions stress placement. We will return to a full analysis of stress in Mandarin and Taiwanese in chapter 3. For the moment, we simply wish to make the point that either of two quite reasonable approaches to defining the environments in which isolation forms of tones occurs is greatly simplified by the perception of neutral toned syllables as toneless.

There is another point worth mentioning. Those morphemes which only occur in neutral tone are the only morphemes in the languages that

do not possess the two forms, isolation and contextual. The language learner is given no evidence whatsoever by which to assign such morphemes one of the eight tones known to exist on other morphemes. If therefore a tone is to be assigned the learner has only two options: either arbitrarily choose one of the eight tones, or assign it a special ninth tone -- special in the sense that it is always accompanied by the diacritic [-stress]. The same problem as with Mandarin then arises: why are there morphemes that alternate between neutral tone and one of the eight lexical tones, but none that alternate within the eight?

Before leaving this section, let us introduce some data from Chaoyang (Zhang 1979) which show rather clearly the way in which a neutral-toned syllable causes the preceding syllable to show up in isolation form, whereas a fully-toned syllable is always preceded by a contextual form. The data involves reduplicated bi-syllabic verbal forms, with the first syllable only reduplicated. If the verb has full tone on both syllables in the simple form, it will also have it on all syllables in the reduplicated form; all non-final syllables undergo isolation to contextual tone sandhi, as expected:

- (1) a. $k'ŋ^{31} \text{ muã}^{55} \rightarrow k'ŋ^{55} \text{ muã}^{55}$ 'to hide things away' 囤瞒
 b. $k'ŋ^{31} \text{ k'ŋ}^{31} \text{ muã}^{55} \rightarrow k'ŋ^{55} \text{ k'ŋ}^{55} \text{ muã}^{55}$
- (2) a. $\text{sau}^{31} \text{ ts'enj}^{33} \text{ k'i}^{31} \rightarrow \text{sau}^{55} \text{ ts'enj}^{33} \text{ k'i}^{31}$ 'clean' 扫清气
 b. $\text{sau}^{31} \text{ sau}^{31} \text{ ts'enj}^{33} \text{ k'i}^{31} \rightarrow \text{sau}^{55} \text{ lau}^{55} \text{ ts'enj}^{33} \text{ k'i}^{31}$

(The segmental change $s \rightarrow l$ observed in 2b is irrelevant here.)

Notice that ... isolation [31] tone is unchanged in final position, but becomes high level [55] elsewhere. (/33/ has identical isolation and

sandhi variants).

Now consider a verb whose final syllable is in neutral tone, which shows up as a low [11].⁵ In the simple form the first syllable is in isolation form, but when it is reduplicated the first syllable of the compound changes to sandhi form; the next syllable, which still precedes a neutral tone, remains unchanged:

- (3) a. siu⁵⁵ k'u³¹ → siu⁵⁵ k'u¹¹ 'swim away' 游去
 b. siu⁵⁵ siu⁵⁵ k'u³¹ → siu¹¹ liu⁵⁵ k'u¹¹
- (4) a. tso³¹³ lou³¹ → tso³¹³ lou¹¹ 'sit down' 坐落
 b. tso³¹³ tso³¹³ lou³¹ → tso³³ lo³¹³ lou¹¹

In all cases the final syllable (which shows up in other morphemes with full tone) shows up as neutral tone. If this is in fact a loss of tone, the preceding syllable is now effectively in final position, and therefore remains in isolation form. When that first syllable is reduplicated, however, the first occurrence does change, from /55/ → [11] or from /313/ → [33]. Note that the final syllable in (3) and (4) is an isolation [31] which shows up as neutral [11]; this is an entirely different change from the contextual change of /31/ to [55] observed in (2).

The only other type of reduplicated adjectives which have the second syllable in isolation form are ones whose final syllable is a morpheme that is always in neutral tone:

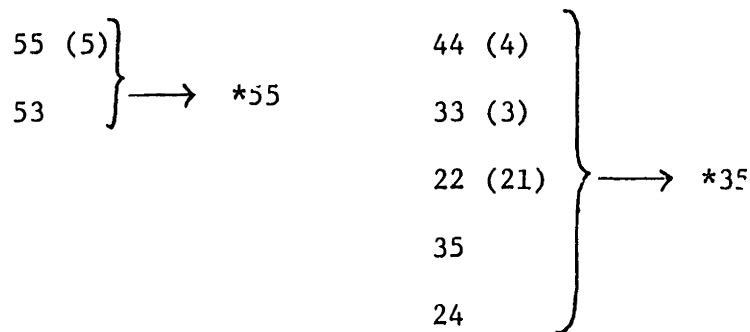
- (5) a. zue⁵⁵ e¹¹ 'To rub/crumple down' 揉下
 b. zue⁵⁵ zue⁵⁵ e¹¹ → zue¹¹ lue⁵⁵ e¹¹

It remains to point out that unlike in Mandarin, neutral toned syllables do not vary phonetically depending on the tone of the preceding syllable. Rather in S.Min they always show a low level pitch. In chapter 5 we will argue that this is due to the insertion of a low tone in foot-final position.

1.3 Deletion of Segments Only

1.3.1 Cantonese Changed Tone

In section 1.1.1 we discussed a process by which the addition of a floating high tone to a morpheme resulted in a phonetic tone change by which all tones which began on a high level became entirely high level, and all tones that began at a lower level acquired a final rise to high. The underlying and derived tones are given again here for ease of reference:



(The changed tone output will be marked with an asterisk for emphasis.)

There is another source of these same derived forms, and the standard literature groups them together as 'changed tone' or 'modified tone' no matter what the source. This second source is through deletion of segmental material, whose tone remains behind. There are thus alternations in the language between forms like the following (examples from Hashimoto (1972) and Whitaker (1955)):

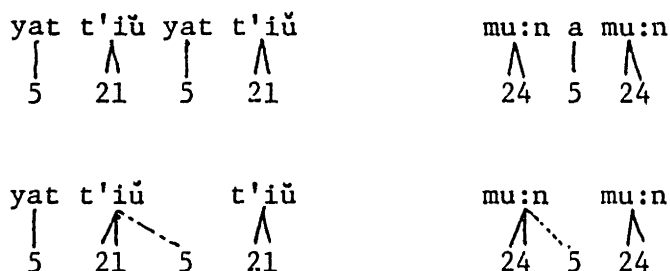
(1) yat⁵ 'one' (in reduplicated measures and verbs, and after the existential verb)

- a. yat⁵ t'iũ²¹ yat⁵ t'iũ²¹
yat⁵ t'iũ^{*35} t'iũ²¹ 'strip by strip' 一條一條
- b. yat⁵ kɔ:⁴⁴ yat⁵ kɔ:⁴⁴
yat⁵ kɔ:^{*35} kɔ:⁴⁴ 'piece by piece' 一個一個
- c. hɔ̃y⁴⁴ yat⁵ hɔ̃y⁴⁴
hɔ̃y^{*35} hɔ̃y⁴⁴ 'go and see' 去一去
- d. yaũ²⁴ yat⁵ kɔ:⁴⁴
yaũ^{*35} kɔ:⁴⁴ 'there's a...' 有一個

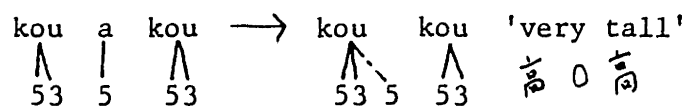
(2) a⁵ 'intensity marker' (in reduplicated adjectives)

- a. hung²¹ a⁵ hung²¹
hung^{*35} hung²¹ 'very red' 紅 〇 紅
- b. mu:n²⁴ a⁵ mu:n²⁴
mu:n^{*35} mu:n²⁴ 'very full' 滿 〇 滿

In both the sets of examples given above the deleted morpheme has a high level tone. If we assume that this tone is left behind by the deletion, then the WFC will require that it attach to the preceding morpheme (pace the question of whether the direction of attachment is language specific, or can be predicted on universal grounds):



If the adjective is one with a high falling /53/ tone, the output of this will be (after simplification) a high level tone:



In fact in these instances it is not possible to state absolutely that the high tone has been left behind, since there is a regular tone sandhi process in Cantonese that simplifies any /53/ that precedes another /53/ (or a /55(5)/) to plain high [55]. Thus the first occurrence of kou⁵³ would surface as [55] either way. However, its form is certainly consistent with the hypothesis proposed here, a hypothesis that is strongly confirmed by the change of tones that begin on non-high levels to rising. In the only example known to me where yat⁵ deletes after /53/ and not before another /53/, the [55] tone indeed results:

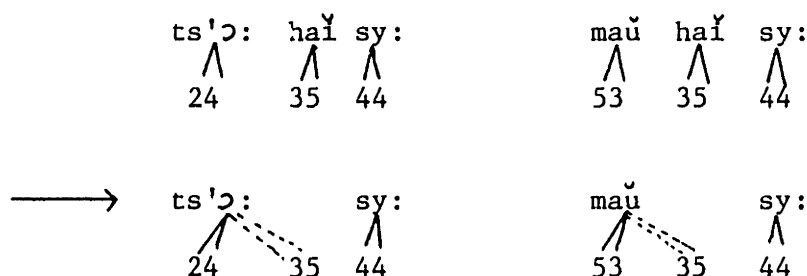
sin⁵³ yat⁵ ts'i⁴⁴ → sin^{*55} ts'i⁴⁴ 'last time'

There are other morphemes that delete, and that have high rising rather than high level tones. For example:

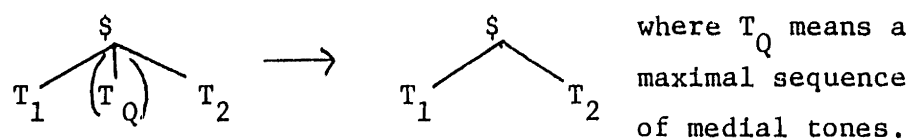
- (3) tsɔ:³⁵ 'perfective' 咗
- sik³ tsɔ:³⁵ → sik^{*35} 'have eaten' 食咗
- faan⁵³ tsɔ:³⁵ lai²¹ tsɔ:³⁵ → faan^{*35} lai^{*35} 'to have returned' 返咗來咗
- kong³⁵ tsɔ:³⁵ lok³³ → kong^{*35} lok³³ 'to have said' 講咗落

- (4) hai³⁵ locative marker 係
 ts'ɔ:²⁴ hai³⁵ sy:⁴⁴ → ts'ɔ:^{*35} sy:⁴⁴ 'sit there' 坐係處
 mau⁵³ hai³⁵ sy:⁴⁴ → mau^{*55} sy:⁴⁴ 'squatting down' 蹲係處
- (5) toũ^{*35} 'to the point that (verb ending)' 到
 fei⁵³ toũ^{*35} p'au³⁵ → fei^{*55} p'au³⁵ 'running away very fast' 飛到跑
 kũa:⁴⁴ toũ^{*35} sy:⁴⁴ → kũa:^{*35} sy:⁴⁴ 'hanging up' 掛到處

These examples provide further evidence for a general process of simplification to the starting point and end point of a tone, with deletion of the intervening tonemes. Consider the following two derivations:

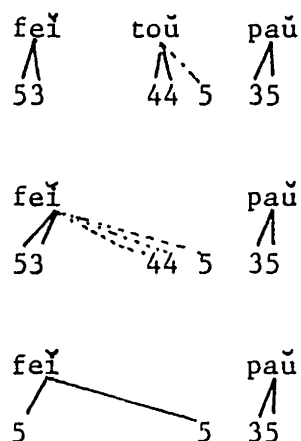


Now the outputs [2435] and [5335] are apparently simplified to [25] and [55] respectively. We can therefore propose a rule of the following rough form:



Assuming that the rule is a late post-cyclic rule that applies only once in a derivation, it may occasionally have to erase even more than two tones. Consider the examples in (5). The morpheme toũ has the underlying tone /44/, but in its use as a verb ending it has

associated with it the floating high /5/ tone justified in section 1.1.1. Suppose that /44/ does indeed consist of two tonemes (although this is not necessarily the case). If it does, then we have the following derivation:



The simplification rule has deleted three tones, leaving only the first and last.

It is noteworthy that all the morphemes that delete optionally and result in changed tones have either a high level, or a high rising tone. This explains why the output always ends on a high level and why there are no alternations between, for example, /24/ and /33/. Other morphemes that occasionally delete include tak⁵ 'can, able', and tsi³⁵, a noun suffix.

We have shown, then, that Cantonese has a process which deletes the segmental layer of certain morphemes, but leaves the tones behind. This is entirely understandable with an autosegmental framework, but constitutes something of a problem for a segmental approach.

1.3.2 Amoy Contraction

C. H. Tung in a footnote in his 1974 article draws our attention

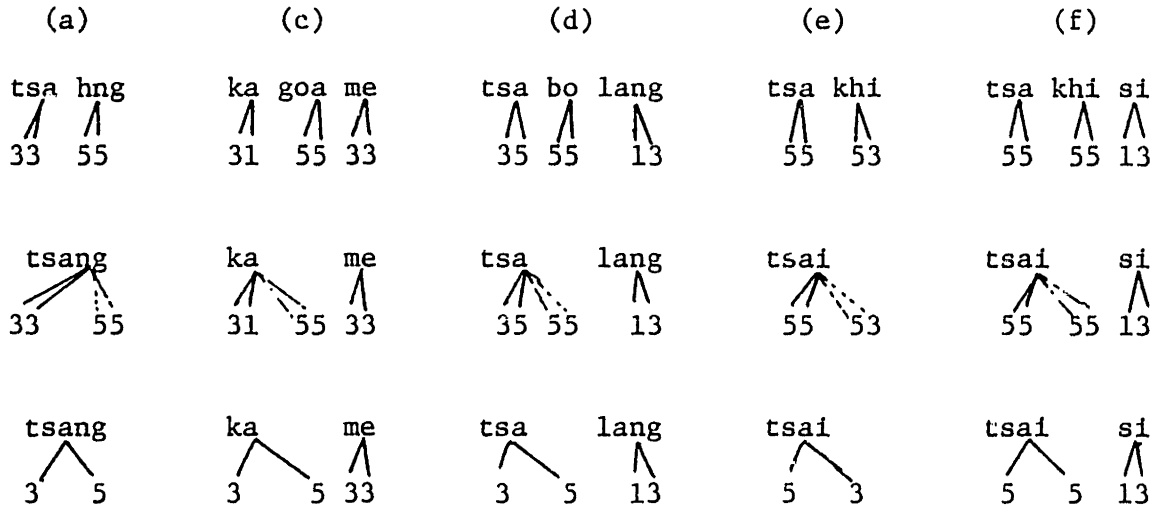
to a phenomenon he calls Syllable Fusion (p. 265):

"Syllable Fusion happens, for example, in kang-pah 'to beat someone' which comes from ka-lang-pah, in hong-me 'to be scolded by someone', which comes from ho-lang-me, in khia-khiai 'to stand up' which comes from khi:-khi-lai etc. In each such case the pitch shapes of the original syllable are retained but shortened in the newly formed syllable, so that sometimes an extraordinary pitch shape results. For example, the pitch shape of chang 'yesterday', which comes from cha-ng (/33-55/) is /35/."

Unfortunately Tung does not give the tones of his examples, so the following are due to R. L. Cheng (personal communication):

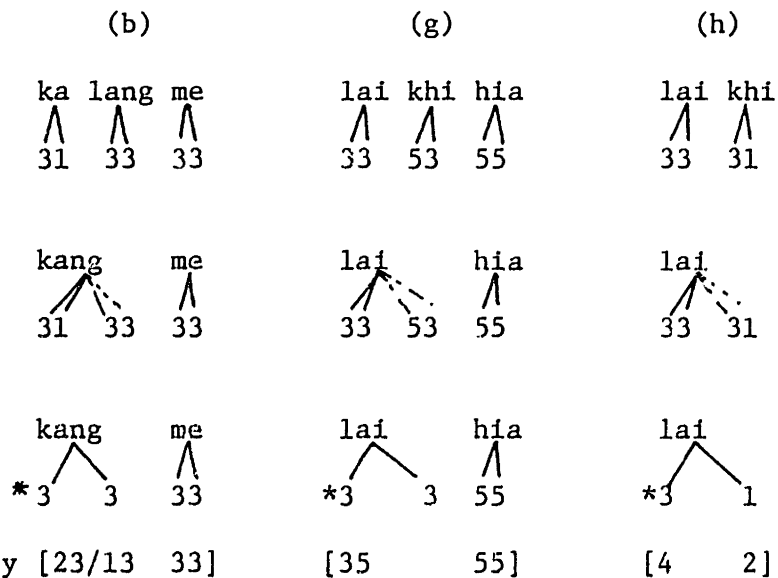
a.	tsa ³³	hng ⁵⁵	tsang ³⁵		
b.	ka ³¹	lang ³³	me ³³	kang ^{23/13}	me ³³
c.	ka ³¹	goa ⁵⁵	me ³³	ka ³⁵	me ³³
d.	tsa ³⁵	bo ⁵⁵	lang ¹³	tsa ³⁵	lang ¹³
e.	tsa ⁵⁵	khi ⁵³	tsai ^{553~54}		
f.	tsa ⁵⁵	khi ⁵⁵	si ¹³	tsai ⁵⁵	si ¹³
g.	lai ³³	khi ⁵³	hia ⁵⁵	lai ³⁵	hia ⁵⁵
h.	lai ³³	khi ³¹	lai ⁴²		

In cases (a, c, d, e, f) the output tone is exactly what would be expected if deletion of the segmentals had left the tones behind, and these had then reassociated with the preceding syllable. Tone simplification then removes any surplus tones:



The only slight difference is that for (e) instead of the post-simplification [53] predicted above, Cheng gives two alternants: [553] and [54]. We will consider the difference small enough to be irrelevant: the output is clearly high falling.

Of the remaining cases, (b, g and h), the following would be the expected derivations:



The first case is given by Cheng as [23] or [13]; [23] is almost level, and may reasonably be taken as equivalent to the predicted [33], a mid-

level tone. [13] however has a definite rise, and it seems that tone simplification for some reason behaves idiosyncratically in this example. Similarly, the second example shows [35] instead of the predicted [33], and it seems that the simplification rule has failed to delete the high /5/ tone. The last example differs from the predicted [31] in being recorded as a slightly higher [42]. However, the prediction is in accord with the facts insofar as both show a falling tone that is lower than the [53] fall seen elsewhere; we may therefore take it that the right prediction is made in this case.

To sum up, the process of Syllable Fusion in Amoy is very simply accounted for in a phonology that allows for the possibility of segmentals being deleted but tone remaining. This tone then reassociates with the preceding syllable, and the resulting sequence is simplified to only the first and last tonemes. In two cases the simplification behaves idiosyncratically in deleting an end toneme instead of a middle one, for reasons that are obscure.

1.4 Deletion of Tone Only

1.4.1 Mandarin Neutral Tone

In section 1.2.1 we introduced Mandarin neutral tone, a term which is used for syllables that have no lexical tone of their own, but rather whose pitch is determined by the tone of the preceding syllable. The alternative, to analyze the neutral tone as a fifth tone, leaves unexplained the asymmetry in the existence of alternations between 'fifth tone' and each of the other four tones, but the absence of alternations between any of the four basic tones themselves. In that section we were concerned with syllables which always show up in neutral

tone, and we argued that these are lexical entries consisting of a segmental component only, with no tonal component. However, it is also the case that there are many many morphemes that show up in some words with full lexical tone, but in others in neutral tone. There are also syntactic conditions under which a few morphemes (object pronouns and verb complements, mainly) can lose their tone. In both these cases we will show that it is necessary to assume that the tones are present underlyingly and are subsequently deleted: an operation that is only possible in a theory that separates out tonal and segmental levels. The argument is based on the third tone sandhi rule that has already been mentioned in connection with the reduplication facts discussed in 1.1.2.

The third tone, written mǎ 馬 has three allotones. In phrase-final position it is low with a final rise [214], and this is the citation form. In non-final position it lacks the final rise, being a plain low [21]. (The initial fall is probably a production phenomenon, since it is hard to start with the voice at the lowest [1] level and easier to take a little time to reach it.) And before another third tone it becomes high rising [35], merging with the underlying second tone, written má 麼. For example, the two phrases below are phonetically indistinguishable (see Wang and Li 1967):

ma ²¹	ma ²¹⁴	→	mai ³⁵	ma ²¹⁴	'buy a horse'	買馬
mai ³⁵	ma ²¹⁴				'bury a horse'	埋馬

Now consider words which are made up of two underlying third tone morphemes (as is apparent from their occurrence in other contexts) the

second of which actually surfaces in neutral tone. The interesting question is what happens to the tone of the first syllable: if it remains a third tone [21] there is no way of telling what the tone of the second syllable was; but if it changes into second tone then a following third tone must have been present at some point in the derivation in order to provide the context for the sandhi rule, and must have been subsequently deleted. In the following words both morphemes can be identified as third tone because they commonly occur as such, but the surface form is second tone-neutral tone, showing that sandhi applied before tone deletion (the examples are from C-C Cheng 1973):

da ²¹ + sao ²¹	→	da ³⁵ sao	'to clean'	打掃
ke ²¹ + yi ²¹	→	ke ³⁵ yi	'may'	可以
xiao ²¹ + jie ²¹	→	xiao ³⁵ jie	'young lady, Miss'	小姐
lao ²¹ + hu ²¹	→	lao ³⁵ hu	'tiger'	老虎

If we look at phrases, such as verb followed by object pronoun, we find the same results:

da ²¹ ni ²¹	→	da ³⁵ ni	'hit you'	打你
gei ²¹ wo ²¹	→	gei ³⁵ wo	'give me, to me'	給我
shui ²¹ li ²¹	→	shui ³⁵ li	'in the water'	水里

(The last example is probably a word rather than a phrase, and the locative li should be treated as a suffix.)

We conclude that it is necessary to assume a rule or rule of tone deletion which leaves the segments unaffected. Such a rule is highly

valued in an autosegmental theory, but impossible in a segmental approach (assuming \pm only, with 0 not allowed).

1.4.2 Shanghai Word Tone

The main part of the Shanghai data that follows is taken from Sherard (1972), supplemented by additional data from Zee and Maddieson (1979). Sherard points out that in polysyllabic words in Shanghai the tone of all except the first syllable is irrelevant to the pitch contour of the word as a whole. For example, the following five words all begin with a syllable in tone one, a high falling tone (about 53). The second syllable carries tones one through five respectively. The pitch contour of all five words is identical, starting high and finishing low (as shown by the superimposed line):

∨ thỳ	+	∨ sàng	thy sàng	'naturally' 天生
∨ kòng	+	┘ sǎ	kong sǎ	'company' 公司
∨ thỳ	+	∨ lhyang	thy lhyang	'dawn' 天亮
∨ thỳ	+	┘ seq	thy seq	'weather' 天色
∨ sàng	+	┘ whəq	sang whəq	'life' 生活

This is just as true if the word is longer than two syllables:

- a) tòng + yhang + nhyng → ~~tong yhang nhyng~~ 'Japanese' 東洋人
- b) mè + thyq + thỳ → ~~me thyq thy~~ 'every day' 每天
- c) sǎ + sang + cǎ → ~~sǎ sang cǎ~~ 'a bastard' 私生子
- d) kòn + nhyng + šyng + chəng → ~~kon nhyng šyng chəng~~ 'worker's housing developments' 工人新村
- e) sǎ + ká + dha + cǎ → ~~sǎ ka dha cǎ~~ 'world war' 世界大戰

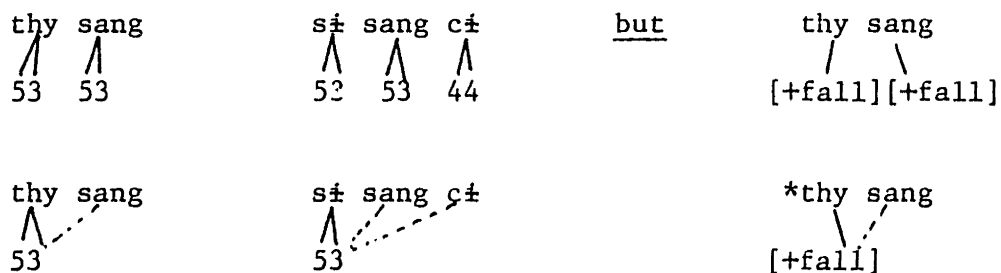
f) čhyng + nhy + kŏ + ho + yhö

青年感化院

~~čhyng nhy kŏ ho yhö~~

'reform school'

Since the tone of the non-initial syllables is irrelevant, it must have been deleted. When this happens the tone of the first syllable spreads out over the now unassociated syllables to its right, and we derive the single pitch contour for all the words above. Zee and Maddieson have shown that contrary to Sherard's description, this contour is exactly as predicted by a level tone theory as opposed to a contour tone theory: the high falling 53 tone spread over two syllables results in high on the first syllable and mid-low on the second, rather than a generalized fall spread over the whole word, or at least [+fall] on the first syllable, as would be expected if a feature [+fall] had spread:



Notice that in the output the tones are in fact associated one-to-one with only [5] associated with the first syllable. This suggests that Shanghai differs from Mandarin in the way tones associate. In Mandarin a sequence of falling tone-neutral tone has the fall on the first syllable, followed by a low:

ˊ mai de / \ 41	'the seller'
----------------------------	--------------

This suggests that in Mandarin all the tones of a morpheme are associated with the segmentals of that morpheme very early on (see section 1.2.1); in Shanghai, on the other hand, it seems that the language specific association rules associated only the leftmost tone with the leftmost syllable (see Clements and Ford: 181-2), and the remaining association lines are inserted by the WFC:

<u>Shanghai</u>	thy sang	versus	<u>Mandarin</u>	mai de
	/			/
	53			41

Shanghai, then, has a process of tone deletion in non-initial position which can be simply stated in an autosegmental framework but makes little sense in any other approach. We will postpone statement of this and the other tone deletion rules until a discussion of stress in chapter 2.

1.4.3 Amoy a⁵³ Suffix

There is a suffix in Taiwanese that has certain rather special properties. The data are from R. L. Cheng (1973). Often called a diminutive suffix, a⁵³ is also used for coordination and as a nominalizer. It forms a very close juncture with the preceding syllable, and triggers resyllabification and voicing phenomena across the syllable boundary. These are discussed in more detail in 2.1.2, and 5.3. All that concerns us here is that the tonal behaviour of words formed by this suffix is distinct from the normal tonal patterns of Taiwanese. The reader will recall that all tones have two alternants, one of which occurs pre-pausally and one of which occurs before another tone. However, before the suffix a⁵³ another set of

alternants occurs, as follows:

	2	3	4	8	7	5	1
Pre-pausal	53	21	21?	54?	33	13	55
Contextual	55	53	54?	21?	21	33	
Before <u>a</u> ⁵³	55		5?	3?	33		

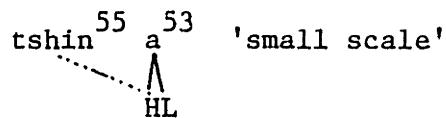
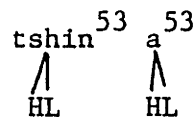
Effectively there are only two variants before this suffix: a high level tone and a mid level tone, each of which occurs also on stopped syllables. Notice now how these relate to the contextual forms, which would presumably be the forms that preceded the more specific pre-a⁵³ varieties. If the contextual form is on a generally high level, be it level [55] (or [54?]) or falling [53], then we find the high level tone. If the contextual form is on a non-high level, whether [33] or [21], then we find the mid level [33]. Notice that it is neither the shape of the contextual tone or its end point that correlates with which variant occurs before a⁵³, because there are falling tones and tones that end on level 3 in both categories. Rather it is the general level of the tone: high or non-high.

We will now have to look forward somewhat to chapter 3, in which we introduce the concept of register. Register in our sense controls the general level of a tone, but its shape is controlled by the tonemes. So a sequence of HL tonemes denotes a falling tone, and if it is in the upper register it is high falling, whereas if it is in the low register it is low falling. Armed with this concept the Amoy facts become easily intelligible.

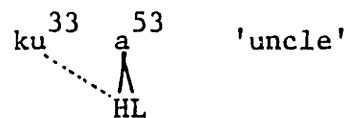
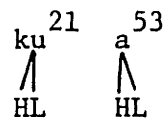
Suppose that the suffix a⁵³ has the special property of triggering

deletion of the tonemes of the preceding syllable, but the register remains behind. Then of course the tone of the suffix will spread backwards onto the preceding syllable to satisfy the WFC, and the newly associated tone will interact with the old register to produce the observed pitch. Below we give two sample derivations, beginning from the contextual forms of the tones:

Upper register



Lower register



When the H tone spreads leftwards to satisfy the WFC, it produces a high tone in the upper register, i.e., the fully high 5 tone on the left, or a high tone in the lower register, i.e., the mid level 3 tone on the right.

Although the concept of register has not yet been fully explained or justified, the reader should be able to see that it is clear that the tones of the first syllable are deleted in some sense, since the contour of that syllable is irrelevant to the final output. For the purposes of this section this is sufficient: there is a rule in the language that deletes tones but leaves the segmentals untouched, and such rules are natural in autosegmental phonology, but not in a segmentally based approach.



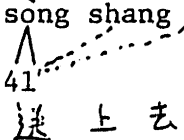
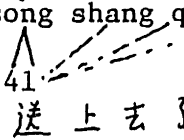
1.5 Spreading

One of the most central phenomena in tonal phonology is the way in which tone spreads out within a domain to all available syllables. The tonal melody, in other words, is realized over varying numbers of syllables and may therefore have rather different surface forms, but all these forms are manifestations of the same underlying melody.

Spreading takes place whenever there are more syllables than tones so that the mapping can no longer be one-to-one, but rather requires that the rightmost toneme be associated with all subsequent syllables. We have already seen ways in which this situation can arise: either there may be more syllables than tones in the lexical entry, as in the example in 1.2.1, or some tones may be deleted by some rule so that certain syllables are left temporarily unassociated, as in the examples in 1.4.

In this brief section we wish to bring together some of those examples, and add a few extra ones to show spreading over greater domains.

Consider Mandarin first. The concept of spreading allows us to treat as tonally identical the forms in (1) a-d, and also the forms in (2) a-d (given that the tone 4 is inserted pre-pausally).

- (1) a. mǎi 'sell' 賣

- b. mǎi de 'seller' 賣的

- c. sòng shàng qu 'send up' 送上去

- d. sòng shàng qu le 'have sent up' 送上去了


- (2) a. māi 'buy' 買
 21 4
- b. māi zhao 'buy successfully' 買着
 21 4
- c. māi zhao le 'have bought successfully' 買着了
 21 4
- d. māi zhao le ma 'have bought successfully?' 買着了嗎?
 21 4

(In fact the later neutral-toned syllables may be lower in pitch as a result of declination.)

A non-autosegmental treatment would need some kind of iterative rule of assimilation (see for example Anderson 1974, Kenstowicz and Kisseberth 1977). When we look at Shanghai spreading the picture is even more complex. Taking the [53] falling tone again, we have suggested that only the leftmost tone is associated by language specific rules; after deletion of the tones of non-initial morphemes spreading will take place as expected. We can therefore unify examples (3) a-e:⁶

- (3) a. thy 'heaven, nature' 天
 53
- b. thy sang 'natural' 天生
 53
- c. tong yhang nhyng 'Japanese' 東洋人
 53
- d. sɿ ka dha cɔ 'world war' 世界大戰
 53
- e. čhyng nhy kɔ ho yhɔ 'reform school' 青年感化院
 53

But in a segmental approach the entire 53 tone is inevitably originally

a property of the initial syllable; it is therefore necessary not only to assimilate left-to-right, but also to subsequently dissimilate the end point of the first syllable so that it surfaces as plain 5, and not 53. All this is unnecessary in an autosegmental framework.

In the preceding five sections we have shown that an autosegmental analysis of various tonal phenomena in Chinese is not only possible, but necessary. We will now continue by discussing two recent proposals that argue for rather different approaches.

1.6 Segmental Tone in Zahao?

In a recent article Andrea Osburne (1979) has argued that Zahao, a Chin language of North Western Burma, must be assumed to have segmental contour tones, and proposes to handle these by means of complex segments. So far we have argued that tone is autosegmental in Asian languages just as it is in African languages, but our arguments have all been drawn from Chinese languages, and it is an interesting question whether tone is necessarily autosegmental, or whether it is simply the unmarked case, and in certain marked languages, such as Zahao, tone might indeed be segmental. Other phenomena which have been handled in an autosegmental framework, such as vowel harmony (Clements 1977), Hebrew morphology (McCarthy 1979), nasality (Goldsmith 1976) and so forth show that a particular feature complex, such as vowel height, may be autosegmental in some languages but not in others: we might therefore assume that the same is true for tone. The difference would then lie in markedness theory: for tone features independence is the unmarked case, whereas for vowel features it is more marked. If it can be shown that a language requires a segmental analysis of tone,

then it is clear that tone, like vowel height, is potentially but not necessarily autosegmental. If on the other hand no such language can be found, then we can only say that so far as current data show tone appears to be always autosegmental, but we must leave open the possibility that future research may turn up the crucial case of a language with segmental tone.

Let us then examine Osburne's arguments for segmental tone in Zahao, which as a Tibeto-Burman language is a member of the greater Sino-Tibetan language family and therefore has some similarities to Chinese languages. The language has three lexical tone shapes: high, low and rising. These are in contrast on sonorant final syllables⁷ but on syllables ending in stops or glottal closure only level high and low are found. There are two sandhi rules which convert underlying rising tones into level tones and which we may write as follows:

(1) LH \longrightarrow H / LH \$ —

(2) LH \longrightarrow L / — \$ H

These feed each other to turn sequences of two rising tones into low tone followed by high tone. It is clear, as Osburne says, that these sandhi can best be handled if the rising tone is represented as a sequence of LH, rather than by a feature [+rise]. Since these rising tones may appear on single vowels it is necessary to find a way of assigning two tones to one vowel, and autosegmental theory can do this without any problem. However, Osburne gives arguments intended to show that an autosegmental approach fails to explain certain other phenomena in Zahao.

The first argument concerns vowel length. Syllables with rising tone are somewhat longer than syllables with level tone. Since vowel length is contrastive in Zahao there are therefore four phonetic degrees of length, as exemplified by:

H	tum	'to play an instrument'	V
LH	tum	'to intend, plan'	V.
L	tuum	'largeness'	VV
LH	tuum	'to be big'	VV.

Osburne says (p. 185):

"Clearly, what is occurring here is that individual vowels are being lengthened to allow time for realization on them of a contour tone, with its two-level (sic) tones of different height in succession. But an autosegmental analysis cannot account for this extra length of vowels with rising tone because it does not map tones onto specific segments."

She therefore concludes that tones must be either segmental, or suprasegmental with its subsequent mapping, and continues to show that they cannot be suprasegmental either, and must therefore be segmental.

Before examining this next argument, let us discuss the first argument from vowel length. The vowel length phenomenon is a late phonetic fact: as far as one can tell from Osburne's article and her thesis (1975). For example, if an underlying rising tone is made level by the sandhi rules (1) and (2), the vowels surface without any

added length. Conversely, if a level toned syllable acquires a contour from intonational effects then it also lengthens somewhat. There seems no reason why lengthening cannot make reference to association lines by lengthening syllables associated with a LH (or any other) tone sequence. When English syllable nuclei are lengthened before voiced stops (bit/bid) there is no suggestion that the trigger must be a feature on the vowel: rather it is a feature of the adjacent segment that causes the lengthening, just as in Zahao it is a feature on the adjacent tier which causes the lengthening: the most ardent proponents of autosegmental phonology have not tried to maintain that there is no interaction between tiers or that rules cannot make crucial reference to features on different tiers. The argument from vowel length, then, is irrelevant to the choice between autosegmental and segmental analyses, as far as one can tell.

Osburne's next argument, against a suprasegmental analysis, involves intonation. In neutral intonation a final low tone is added to high and rising tones. So $H + L \longrightarrow HL$ and $LH + L \longrightarrow LHL$. She argues that this must happen after the lexical tones have already been mapped onto the segments since it creates new contours which do not exist underlyingly, and that since it also results in vowel lengthening because it creates a contour tone it must be actually mapped, rather than just associated. But if intonation is therefore suprasegmental, lexical tone cannot be, because two separate mappings would be required. This argument appears to me to be entirely inconsistent. It rests on the assumption that if a process creates unusual contours, that it must do so by unusual means. On the contrary, it is a strong argument for the explanatory power of

any theory if the same mechanisms that account for one set of alternations also account for an entirely different set of alternations. There seems no reason whatsoever to assume two sets of mappings (or associations) since a single left to right mapping (or association) would also capture the facts. And even if two sets were required, normal notions of cyclicity would make that entirely unsurprising. This argument, then, also fails to show that tone in Zahao must be treated as segmental.

The next three arguments are supplementary. Osburne's third argument comes from the absence of tonal melodies like those found in African languages: one might expect that the same tones that show up on monosyllables as L, H, LH might surface on bisyllables as L.L, H.H, L.H, but in fact each syllable may bear one of the three lexical tones, and there is no evidence of melodic spreading. This argument fails on two grounds. First, the absence of melodies is not of itself an argument: the presence of melodies is. Secondly, Osburne (1975: 55) states that the syllable and the morpheme are essentially coextensive in Zahao, as in most S.E. Asian languages. In that case the absence of melodies over sequences of syllables is the result of the fact that lexical tones are properties of morphemes, and therefore, (in Zahao) of syllables.

The fourth argument looks quite convincing, since it deals with elision. One of the clearest predictions of autosegmental theory is that deletions on one tier will not necessarily result in deletions on another tier. In particular, then, segments may be deleted while tones are left behind. These tones will then re-associate to other segmental material.

This notion of melody stability is a cornerstone of autosegmental theory, and counter-examples would constitute a strong argument in favour of the segmental nature of tone in a language. Osburne gives the following examples of the elision of the vowel in the adverbial particle ?in:

L L H LH	→	L L H	L L HLH	
zanrangte ?in		zamrangteen	(not* zamrangteen)	'quickly'
H LH	→	H	HLH	
na ?in		naan	(not *naan)	'but'

The particle has a rising tone, and when the two syllables are elided into one the result is not a complex tone made up of the underlying sequence, but rather the rising tone itself is lost, as is the initial glottal. The vowel then assimilates to the preceding vowel. There are two reasons why it is not at all clear that this constitutes an argument in favour of a segmental analysis of Zahao tone. Firstly, what is actually deleted is the tone and the glottal stop, not the vowel. Both these are laryngeal phenomena, and as such this may in fact be an argument in favour of an independent laryngeal tier, on which both tone and glottalization are represented, rather than against it. The deletion is then of the information on this laryngeal tier, while the segmental tier remains untouched. Subsequently the two adjacent vowels assimilate. This is shown below:

na in	→	na in	→	naan
\		\		\
H ?LH		H		H

Secondly, it is not clear from Osburne's article how general a

phenomenon this elision is, but in her thesis (1975: 59-60) she states specifically that this adverbial particle is the only morpheme that elides in normal speech. It is therefore apparently a morphological rather than a phonological rule, and as such would not necessarily be expected to obey the predictions of the stability hypothesis.

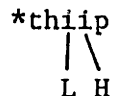
Osburne's last argument concerns the segmental constraints on tones. As we mentioned earlier, rising tones are only found on unstopped syllables. Osburne's point is that these contextual constraints are more elegantly stated if tones are features on segments than by making reference to two tiers and the permitted associations between those tiers. This is probably the strongest argument, since autosegmental phonology predicts that any number of tones can be associated with any tone-bearing unit unless special (and therefore expensive) restrictions are stated language specifically. Even this, however, does not actually require that tone be analyzed as segmental: it simply suggests that an extra condition must be put on the system restricting multiple attachments under most circumstances. Such conditions are very common (e.g., Lithuanian) perhaps even the unmarked case. In all cases of rising toned syllables except one there is a sequence of sonorants present. If then we take it that the sonorant in the rime is the tone bearing unit in Zahaó and association is one-to-one, rising tones will involve sequences of the following kind:

$\begin{array}{cc} t & u & m \\ & & \\ L & H & \end{array}$	or	$\begin{array}{ccc} t & u & u & m \\ & & \vee & \\ & L & H & \end{array}$
---	----	---

In syllables with only one sonorant we have multiple association:



but these syllables are interesting in that their tone is always simplified to a level high tone utterance internally. This suggests that multiple associations are avoided in Zahao except utterance finally, where they can arise in two ways: either from underlying sequences of the kind shown here, or by addition of an extra intonational boundary tone. Under the latter circumstances even stopped syllables may acquire rising tone. It still remains to account for the fact that there is clearly a morpheme structure condition that differentiates the following two configurations:



allowing the first but not the others. If it is the case that the stopped syllables all involve glottal closure (i.e., that thiip as well as qaq has glottal closure) then the constraint is arguably not an inter-tier constraint, but rather to be stated on a single laryngeal tier. The absence of syllables of the third type above, *thiip shows that it is not sufficient to block multiple associations in the presence of glottal closure: rather the sequence /L H [+glottal]/ is disallowed.

To conclude, the data from Zahao do not seem to argue in favour of a segmental analysis of tone, and could easily be accounted for within an autosegmental analysis. We can therefore still maintain the hypothesis that tone is normally autosegmental, and that the child will

so analyze it unless forced to do otherwise. In Zahao the child has no reason to suppose that tone is anything but autosegmental.

Before leaving this question, it is interesting to take a brief look at a language that Leben (1971 and 1973) analyzed as having segmental tone. The argument, like Osburne's, was based on an apparent relationship between vowel length and tone, and the language was Thai. Leben's data comes from Henderson (1949).

Leben pointed out that although Thai has two contour tones (one rise and one fall) in addition to three level tones, in the combinative speech style these are simplified to mid tones on the first syllable, and at the same time the vowel is shortened:

(1)	$\begin{array}{c} \text{thi:} \\ \wedge \\ \text{HL} \end{array}$	$\begin{array}{c} \text{nai} \\ \wedge \\ \text{LH} \end{array}$	→	$\begin{array}{c} \text{thi}' \\ \\ \text{M} \end{array}$	$\begin{array}{c} \text{nai} \\ \wedge \\ \text{LH} \end{array}$	'where?'
	$\begin{array}{c} \text{si:} \\ \wedge \\ \text{LH} \end{array}$	$\begin{array}{c} \text{kha:u} \\ \wedge \\ \text{LH} \end{array}$	→	$\begin{array}{c} \text{si}' \\ \\ \text{M} \end{array}$	$\begin{array}{c} \text{kha:u} \\ \wedge \\ \text{LH} \end{array}$	'white'
	$\begin{array}{c} \text{sa:u} \\ \wedge \\ \text{LH} \end{array}$	$\begin{array}{c} \text{sa:u} \\ \wedge \\ \text{LH} \end{array}$	→	$\begin{array}{c} \text{sau} \\ \\ \text{M} \end{array}$	$\begin{array}{c} \text{sa:u} \\ \wedge \\ \text{LH} \end{array}$	'young girls'
	$\begin{array}{c} \text{wa:n} \\ \wedge \\ \text{HL} \end{array}$	$\begin{array}{c} \text{wa:n} \\ \wedge \\ \text{HL} \end{array}$	→	$\begin{array}{c} \text{wan} \\ \\ \text{M} \end{array}$	$\begin{array}{c} \text{wa:n} \\ \wedge \\ \text{HL} \end{array}$	'at your leisure'

Leben then points out that if the first syllable bears a level tone the vowel shortening still takes place, but there is no accompanying change in tone:

(2)	$\begin{array}{c} \text{na:m} \\ \\ \text{H} \end{array}$	$\begin{array}{c} \text{tɕha:} \\ \\ \text{M} \end{array}$	→	$\begin{array}{c} \text{nam} \\ \\ \text{H} \end{array}$	$\begin{array}{c} \text{tɕha:} \\ \\ \text{M} \end{array}$	'tea'
-----	---	--	---	--	--	-------

(Leben's examples show only high toned cases, but Henderson also gives a low toned case:

(3) ja:n nan → jaŋ ɲan 'that way, like that')
 | | | |
 L H L H

He therefore suggests that these facts follow if there is a rule of the following form:

VV → V

and a convention that the phonetic result of simplifying HL or LH is a 'compromise' mid tone. The failure of the tones to simplify when the vowel shortening rule does not apply is thus explained:

(4) thau rai → thau rai 'how much'
 / | / |
 HL M HL L

 tɔŋ ka:n → tɔŋ ka:n 'want'
 / | / |
 HL M HL L

There are a number of problems with this argument. Firstly, Leben is forced to the following statement:

"tone (is)...a segmental feature that may appear on vowels and on the coda of short syllables ending in a voiced segment."

This is a very odd distribution of tone, with its presence on the coda conditioned by the complexity of the nucleus. One would expect tone to be either limited to the nucleus, or free to appear anywhere in the rhyme, but neither can be the case for Leben. If it were limited to the nucleus, then no contour tones should be possible in words like [tɔŋ], with a simple nucleus. And if it were free to appear on the entire rhyme

then words like $\left[\begin{array}{c} *wa:n \\ HLH \end{array} \right]$ should be found with three tonemes. One other possibility should be considered: one tone is assigned per nucleus, and one per coda. But if tone is a feature on segments this kind of assignment is hard to deal with, and quite against the spirit of segmental tone if not the letter.

The second problem is that if tone is truly segmental one would expect it to delete when the vowel deletes: it is not at all obvious why its effects should remain behind. The facts could thus be turned around to argue against a segmental representation of tone, rather than in favour of it.

When one examines Henderson's data more closely there are other problems for a segmental analysis. For example, in sentence intonation additional tone patterns can arise including contour tones on short vowels followed by glottal stop or zero:

(5) ba(?) la(?)
 \wedge \wedge
 HL HL

(Henderson's Sentence Tone D: 45-6, and 44: fn. 7.)

Secondly, also in sentence intonation, vowel shortening may optionally take place without any tone simplification (p. 47):

(6) kha(:)u -ma: si: Do come in!
 \wedge | |
 H L M H

Thirdly, Henderson gives cases of tone change accompanying vowel shortening with underlyingly level tones:


(7)	$\begin{array}{c} \text{ja:ŋ} \\ \\ \text{L} \end{array}$	$\begin{array}{c} \text{rai} \\ \\ \text{M} \end{array}$	→	$\begin{array}{c} \text{jaŋ} \\ \\ \text{M} \end{array}$	$\begin{array}{c} \text{ŋai} \\ \\ \text{M} \end{array}$	'how'
	$\begin{array}{c} \text{ja:ŋ} \\ \\ \text{L} \end{array}$	$\begin{array}{c} \text{ni:} \\ \\ \text{H} \end{array}$	→	$\begin{array}{c} \text{jaŋ} \\ \\ \text{M} \end{array}$	$\begin{array}{c} \text{ŋi:} \\ \\ \text{H} \end{array}$	'this way, like this'

The last example forms a minimal pair with (3), which suggests that the tone change may be morphological rather than phonological. Further support for this proposal comes from the fact that Henderson also discusses the neutral tone (unstressed syllables) and points out that (p. 37, fn. 27):

"The actual pitch of the neutral tone may vary according to context, but is most commonly mid level."

Because of the gaps in Henderson's data (for our purposes) it is not possible to reach a firm conclusion, but I suggest that the tone simplification may be stress-conditioned (not by complete absence of stress, like the neutral tone, but rather by secondary as opposed to main stress). This is supported by a remark in another paper by Henderson (1967) in which she states that the stress in the form

$\begin{array}{c} \text{thi:} \\ \wedge \\ \text{HL} \end{array}$	$\begin{array}{c} \text{nai} \\ \wedge \\ \text{LH} \end{array}$	→	$\begin{array}{c} \text{thi}' \\ \\ \text{M} \end{array}$	$\begin{array}{c} \text{nai} \\ \wedge \\ \text{LH} \end{array}$	'where'
					(given in (1) above)

falls on the second syllable (she writes ). She then points out that this makes the source of the tone change unclear: it may be a sandhi rule, but it may rather be stress-conditioned. Vowel shortening, on the other hand, happens automatically in non-final position by a rule like Leben's. The tone, however, remains behind, as evidenced by examples like (6) (which may undergo emphatic lengthening optionally).

Even if this analysis should turn out to be wrong, it is clear that there is no overwhelming argument for taking tone to be segmental in Thai. Elsewhere (1973: 34) Leben mentions that in one type of Thai word game tones move with their vowels:

$$\begin{array}{ccc} \text{kon} & \text{jaj} & \longrightarrow & \text{kaj} & \text{jon} & \text{'big bottom'} \\ \wedge & | & & | & \wedge & \\ \text{HL} & \text{L} & & \text{L} & \text{HL} & \end{array}$$

and suggests that this is to be expected in a language with segmental tone as opposed to one with suprasegmental tone. However, there are other Thai word games in which the tone does not move with the vowel (see section 0.1.2) and these are inexplicable if tone is segmental. On the other hand if it is suprasegmental both types are possible, one applying before tone association (or mapping, for Leben) and one after. Leben himself points this out (1973: 162):


"...even in a language for which a suprasegmental level of representation is established, there can exist rules which treat tone as a segmental feature. Such rules, however, must be ordered after the rule or rules which map suprasegmentals onto individual segments."

(This same distinction arises in Mandarin reduplication, as we will see in section 1.1.2.)

1.7 A Dynamic Tone Theory

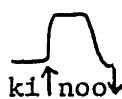
1.7.1 A Brief Summary of Dynamic Tone Theory

In her thesis Mary Clark (1978) proposes a theory of tone that

is very different from either the segmental approach as exemplified by Woo (1969) or any of the better known suprasegmental approaches, such as Leben (1973) and Goldsmith (1976). In Clark's theory the primitives of tone are neither level tones, nor contour features; rather they are dynamic tone markers written \uparrow and \downarrow , which are instructions to change pitch. These markers are interspersed among the segmentals of the language and carry information about the relative pitch of two adjacent segmental units. For example, suppose that in one language they appear between syllables. Then a sequence of $\$ \uparrow \$$ would denote a pitch contour that began at an unmarked level and rose so that the second syllable had high pitch. It is therefore equivalent to the sequence LH. In some languages these markers may appear between moras as well as between syllables, and this is one way of characterizing a rising tone. For example, the syllable made up of $\mu \uparrow \mu$ would have the pitch contour . Tone markers may appear before and/or after syllables and moras, and the exact phonetic realization rules will vary from language to language. Suppose some language allows final tone markers. Then a sequence $\$ \uparrow$ may either denote a low tone, so that the \uparrow is realized as the transition into the next syllable, or it may denote a rise on the syllable preceding it. It may be the case that a marker is realized in one way under some circumstances, and another way under other circumstances in the same language. For example, Clark (p. 27) gives the following examples of the way the sequence $\dots \uparrow \dots \downarrow \dots$ is realized in Japanese:



'man'



'yesterday'



'heart'

The marker ↓ is realized on the following syllable if possible, but in absolute final position it is realized on the preceding syllable as a falling glide. Clark points out that in this way her system is able to capture the equivalency between a sequence of LH on two syllables and a LH rise on a single syllable -- thereby circumventing one of the shortcomings of Wang's system.

Clark argues in section seven of chapter one (pp. 36-47) that her theory has several advantages over autosegmental theory. The first two arguments concern indeterminacy: she claims that autosegmental theory does not always allow one to distinguish unambiguously between accentual and non-accentual systems, and that it provides too rich an inventory of possible rules.

The first of these observations, ambiguity between accentual and non-accentual systems, is probably correct, but not necessarily a disadvantage. It is highly likely, as Clements and Ford (1979: 199-202) have argued, that the correct distinction is rather between free accent and fixed accent systems, and that one can change into the other diachronically. If this is right, ambiguous cases may be rightly ambiguous in that they represent languages in transition from one type of accent to the other. The second criticism -- that there are too many possible solutions to a given phonological problem because of the rich inventory of rules -- is valid, but hardly a reason for dismissing the theory. It is undoubtedly true that more work is needed on constraining the theory and capturing notions such as possible rule, highly valued rule, etc., but the same is true for Clark's theory.

Clark argues (p. 39) that autosegmental theory intrinsically allows more rule types than dynamic tone theory because it has three layers of

tonal structure to play with: starred elements, tone melodies, and association lines. Dynamic tone theory on the other hand has only tone markers. The types of autosegmental rules she mentions are:

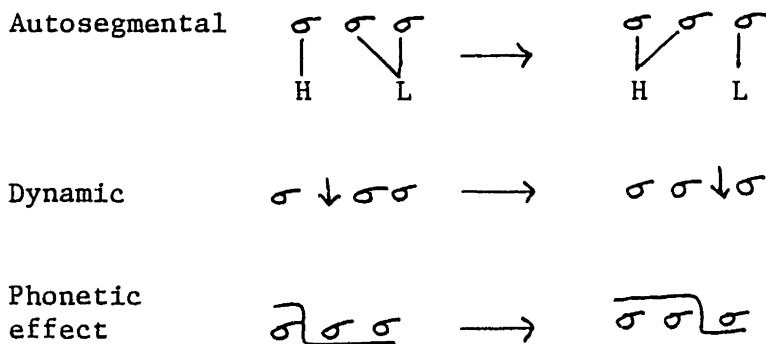
- (i) rules inserting, deleting or moving a *
- (ii) rules inserting, deleting or changing tones and phonemes
- (iii) rules associating or de-associating tones and phonemes
- (iv) rules transposing two tone segments

Let us compare these to dynamic tone rules. It is true that no * is used in Clark's theory, but if one could maintain the distinction between morphological rules, which act on *, and phonological rules, which do not, the power of the theory is not seriously increased, and indeed if it is possible to further state that only * rules can be morphologically conditioned, discussion reduces to the old problem of the 'best theory' (Postal 1972). Clark's position would essentially be the same as Postal's, I assume: a homogeneous theory is a priori to be preferred in the absence of evidence to the contrary. I would take the opposing view that the more articulated theory made up of several distinct sub-systems, each of which is highly constrained, may be at least as suitable. Specifically, if * rules can be shown to be identical with morphological rules and tone rules identical with phonological rules, then there is no loss of explanatory power whatsoever. Of course, it remains to be shown that this is the case, but it is the obvious direction to move in in searching for a more constrained theory of tone.

Clark's type (ii), rules inserting, deleting, and changing a tone segment, has its formal correlate in her theory in rules inserting,

deleting and changing pitch change markers. She admits rules of both the first two types, but rules changing tone markers are restricted to changes in the size of markers (from full size \uparrow to half size \uparrow) and she excludes rules changing \uparrow to \downarrow and vice-versa (p. 171). Notice what the effect of such rules would be. A representation like $\sigma\downarrow\sigma$ would be changed to $\sigma\uparrow\sigma$, and this would amount to changing falling to rising tones. The exclusion of this type of change is therefore exactly equivalent in effect to the exclusion of tonal metathesis in an autosegmental framework, and Clark notes that such rules (type (iv) in the list above) are exactly the type of rule that Haraguchi (1975) does not make use of. Notice however that there is nothing within dynamic tone theory that would make such rules impossible (except their exclusion by fiat) whereas in autosegmental theory they require the use of transformational rules, a power that is not generally necessary in the tonology and that can probably be excluded altogether.

Clark's third autosegmental rule type makes changes in association lines. The formal equivalent in dynamic tone theory are the movement rules that move tone markers across syllables. For example, the following are equivalent:



Although these two rule types are the nearest equivalents, it is

important to point out that their effects are not always the same by any means: for example, moving a single pitch change marker one syllable right or left can change a low tone to a rising or falling tone:

$$\begin{array}{ccc} \sigma\sigma\uparrow & \longrightarrow & \sigma\uparrow\sigma \quad \text{↗} \\ \downarrow\sigma\sigma & \longrightarrow & \sigma\downarrow\sigma \quad \text{↘} \end{array}$$

The point at issue here is the general power of the theory, however. If the existence of rules that change association lines in some way is a mechanism available only to autosegmental theory, it is equally true that rules which move tone markers over syllables are a mechanism available only to dynamic tone theory.

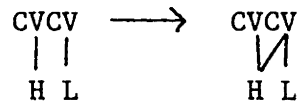
The fourth rule type, transposition rules, has been discussed above in connection with metathesis.

Before leaving this section it is instructive to look at Clark's illustration of the diversity of options for explaining a single phenomenon within autosegmental theory. The example is hypothetical (but apparently based on real facts in Japanese). Suppose that in a language with a basic HL melody, the tone association rule of standard Japanese, and a tone simplification rule of the form:

$$(7.2-1) \quad T \longrightarrow \emptyset / \begin{array}{c} V \\ \swarrow \quad \searrow \\ T \quad \underline{\quad} \end{array}$$

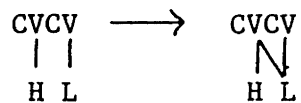
it were to turn out that in certain circumstances bisyllabic nouns with accented first syllables had a high level tone contour instead of the expected HL. Clark points out that there are seven possible analyses of these facts in the autosegmental theory:

- a) The * accent moves from the first to the second syllable
- b) The * accent is deleted
- c) There is a H-spreading rule:



Rule (7.2-1) then deletes the L tone.

- d) There is a L-spreading rule:



Rule (7.2-1) then deletes the L tone, and the H spreads by the WFC.

- e) There is a L deletion rule, and the H spreads by the WFC
- f) There is a L \longrightarrow H assimilation rule
- g) There is a H tone insertion rule, and (7.2-1) then deletes the L:



In fact this diversity is misleading: there are a number of grounds for choosing between the various possibilities, both empirically and theory-internally. Let us narrow the field down. The first two possibilities involve * , and would therefore be expected in a morphological rule but not a phonological rule. It turns out that the word class in which this phenomenon is found is a syntactic class (Clark 1978b: 93), so in fact one of these (a-b) is to be preferred. The choice between them can be made on empirical grounds if they are ever followed by affixes, since (a) predicts low tone on the affix, while (b) predicts H tone. If this evidence is lacking, it would be desirable if the theory forced us to one alternative, presumably (b).

Suppose that this hypothetical situation were found in a phonological context. In that case (a-b) would be excluded on principled grounds, and a choice must be made between (c-g).

One kind of evidence to force a choice would be the environment for the rule. If tone rules are local, as seems likely, a rule like (e) which deletes the second tone should be conditioned by the right-hand environment not the left-hand environment, and the same is true for (f). If then the environment were to the left these two could be excluded. Another kind of evidence would be the subsequent behaviour of these words. The output of rules (c-e) is $\begin{array}{c} \text{CVCV} \\ | \\ \text{H} \end{array}$, whereas (f-g) produce $\begin{array}{c} \text{CVCV} \\ | \quad | \\ \text{H} \quad \text{H} \end{array}$. These would be phonetically non-distinct, but phonologically distinct, and might condition or undergo later rules differently. Option (d) can probably be excluded on different grounds. Notice the effect of rule (7.2-1) in this case:

$$\begin{array}{c} \text{CVCV} \\ | \quad | \\ \text{H} \quad \text{L} \end{array} \xrightarrow{7.2-1} \begin{array}{c} \text{CVCV} \\ | \\ \text{H} \end{array} \xrightarrow{\text{WFC}} \begin{array}{c} \text{CVCV} \\ | \quad / \\ \text{H} \end{array}$$

Simplification rules of this kind do not usually delete tones that are associated with other syllables also. Rather they are low level phonetic rules that simplify contours on single vowels, and as such only delete the tone entirely if it is associated with no other syllable. In this instance, then, a proper formulation of 7.2-1 would erase the association line but leave the tone intact, which does not produce the desired output.

The final criterion one might apply is not empirical, but rather a reasonable constraint on the power of the theory. Lexical entries

include tones and phonemes but not, usually, association lines. Changes in tones and phonemes might therefore be considered more expensive in that they could involve a loss of information, whereas association lines carry no such lexical information. If this is right, rule (c) is clearly to be preferred over rules (e-g), and if (d) is already excluded the possibilities are reduced to one.

Compare this to the dynamic tone theory, which allows exactly two accounts of the above facts:

(a) delete the underlying ↓ in such words:

$$CV \downarrow CV \longrightarrow CVCV$$

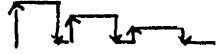
(b) move the ↓ to the end of the word:

$$CV \downarrow CV \longrightarrow CVCV\downarrow$$

These make different predictions about the tone of following affixes, as do (a) and (b) in autosegmental theory. If there is no such evidence available, the choice must be made on other grounds, and Clark must therefore also face the problem of restricting the number of rule types available.

The foregoing discussion has been intended to give no more than a suggestion of the kinds of ways in which the power of the theory can be restricted. However, I hope to have shown that the task is by no means as hopeless as Clark implies, and that in fact dynamic tone theory must also deal with this problem.

Clark's last argument maintains that the dynamic tone theory allows for a particularly elegant treatment of downdrift in languages as disparate as Igbo and Japanese. She introduces a convention that

makes "a pitch rise at any point within the breath group a little smaller than a pitch drop would be at that same point...The result will be an overall lowering of the pitch register of the voice throughout the breath group." She goes on to show that although it is possible to account for some cases in autosegmental theory by means of a convention which lowers the high and low pitch registers whenever unlike tones succeed one another, there are some facts that require addition of more complex conditions that lack any explanatory power. While Clark's analysis of downdrift is very attractive for languages like Igbo that downdrift H and L tones, it needs modifications for a language that downdrifts H only. The effect can be pictured as 

and requires one of two statements: "a pitch change is always the same size as the preceding rise, which is always smaller than the last rise" or "each pitch change is smaller by a constant amount than the last change in the same direction." In an autosegmental theory the registers can be lowered in both cases, the difference being only whether lowering involves the low register or not.

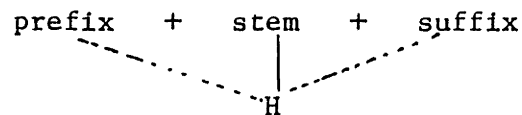
There are a number of both general and specific problems with Clark's approach, some of which stem from the choice of dynamic rather than level primitives, and some of which stem from the decision to intersperse the markers among the segmentals. To a certain extent that decision is inevitable, since the location of the pitch marker is crucial. Clark allows languages to distinguish $\mu \uparrow \mu \downarrow$ from $\uparrow \mu \downarrow \mu$ (see Chaozhou, p. 145) for example, so that if the markers formed an independent tier associated with the segmental tier the segmental tier would have to include some sort of abstract location markers to condition the position of association. Chao and Young

(1979) have made an interesting attempt to work out an approach of this kind that avoids many of the problems of Clark's approach, in particular those discussed in 1.7.3.

1.7.2 Problems Stemming from the Choice of Dynamic Primitives

(a) Direction of Spreading

In an autosegmental theory all segments must be associated with some tone. Therefore if there are more syllables than tones the tones will spread onto all unassociated syllables. Suppose there is a stem with associated H tone, and toneless prefixes and affixes are added. The H tone will then spread in both directions:



In Clark's dynamic tone theory the prediction is different. Languages are claimed to be of two types, those which use initial tone markers and those which use final tone markers. The prediction would be that spreading is unidirectional, and that the direction of spreading is different in languages of the two types. So a language with final tone markers (motivated on independent grounds) like Igbo, represents high tone as follows:

\$↓ \$ \$↓ or \$ \$ \$↓ (p. 88)

This marker denotes a change of direction so if a toneless syllable is added to the end of such a word, giving:

\$ \$ ↓ \$



the pitch contour will drop down on the final syllable (to the unmarked mid-level rather than to a low tone). That is, there will be dissimilation rather than spreading. On the other hand if a toneless prefix is added the tone will effectively spread:

§ § § ↓ ———

In a language like Mende, which has initial tone markers, the situation will be reversed, so that tone will spread rightwards onto toneless affixes, but not leftward onto toneless prefixes. If there were indeed such a correlation it would provide strong support for Clark's theory, but it does not seem to be correct. For example, Igbo verbs can be suffixed. When a suffix is added to a low-toned verb stem, which has a final ↑ marker, the theory predicts that the suffixes should show up with high or mid tone:

verb ↑ + suffix ———

But in fact they show up with low tone, exactly as predicted by autosegmental theory:

verb + suffix ———
|
L



Clark is therefore forced to postulate a condition requiring all words to have final tone markers, and provides for the copying of the verbal marker ↑ to the end of the word (p. 297), giving:

verb ↑ + suffix ↑

The first marker is then deleted by a general rule of like-marker

deletion (see (b) below). None of this is necessary in autosegmental theory, in which the WFC will automatically produce the right result.

It is interesting to note in this context that Igbo does have a couple of verbal prefixes, and they exhibit the following behaviour: the negative a always has H tone, but the affirmative a has the opposite tone to the verb stem (although it has usually been analyzed as underlyingly low toned). Notice that the polarizing behaviour of the affirmative a is exactly what one would expect in dynamic tone theory if the prefix was toneless, and Igbo verb stems had initial tone markers instead of the final markers assumed by Clark:

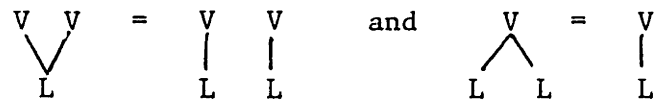
a	+	↓	verb stem	
a	+	↑	verb stem	

But as her current analysis of Igbo stands verb stems have final tone markers only, in which case suffixes should be polarizing, not prefixes. Clark (p.c.) has advanced the possibility that Igbo has word-final tone markers, but stem-initial markers. This would cause considerable complication in the tonology, but would explain the behaviour of affixes.

(b) Like Tone Marker Deletion

One of the main problems of the dynamic tone marker approach is what happens when sequences of syllables with the same tone arise. Take a language like Mende in which level tones are marked by tone markers at the beginning of words and final markers, if present, are realized as glides on the preceding syllable:

In a level tone theory sequences of this kind are portrayed as sequences of H H or L L etc., and nothing particular needs to be said to account for the surface pitch contour. The nearest equivalent is the possible existence of low level phonetic rules collapsing



Furthermore, the fact that the directionality of the like-tone marker deletion rule correlates with the positioning of the tone markers suggests that a generalization is being missed somewhere; two languages which exhibit the same entirely unremarkable behaviour in allowing sequences of high or low tones to surface as such are supposed to differ in two ways, both in the positioning of the markers and in which deletes. Clark's case might be more convincing if there were clear cases of languages which had a different combination of positioning deletion. Presumably such a language would look like this:

$$\begin{array}{l} \uparrow\sigma \uparrow\sigma \rightarrow \sigma\uparrow\sigma \quad \text{i.e.} \quad H H \longrightarrow M H \\ \downarrow\sigma \downarrow\sigma \rightarrow \sigma\downarrow\sigma \quad \text{i.e.} \quad L L \longrightarrow M L \end{array}$$

assuming that syllables with no preceding marker would be realized on a mid-level.

(c) The Inventory Problem

Although Clark only has two tone markers \uparrow and \downarrow (in fact she mentions the possibility of adding markers of different sizes) she is able to describe an enormous number of contrasts. The following

are the seven possibilities in a language with markers at syllable boundaries only:

$$\begin{array}{ccc} \uparrow\sigma & \sigma\downarrow & \uparrow\sigma\downarrow \\ \downarrow\sigma & \sigma\uparrow & \downarrow\sigma\uparrow \\ \sigma & & \end{array}$$

The phonetic realization of these depends on the language. For example $\uparrow\mu\downarrow$ is a high level tone in Chaozhou (p. 145) and so is $\uparrow\mu$. $\uparrow\sigma$, (like $\uparrow\mu$), is a high level tone in Mende, and $\sigma\downarrow$ is in Igbo. There are therefore three possible ways of representing a high level tone, more than one of which may be used in a single language (see Chaozhou). In addition to the possibility of multiple representations for one phonetic form there is ambiguity in the opposite direction too, since one representation may have multiple phonetic realizations. For example, $\downarrow\sigma\uparrow$ is a rising tone in Mende, but $\downarrow\mu\uparrow$ is a low level tone in Chaozhou.

Furthermore, once inter-moraic tone markers enter the picture there are more possibilities for contour tones. All the following four are possible representations of rising tones, and to them must be added the Mende-type possibility involving the syllable only, for a total of five:

$$\begin{array}{ccc} \mu\uparrow\mu & \downarrow\mu\uparrow\mu & \downarrow\sigma\uparrow \\ \mu\uparrow\mu\downarrow & \downarrow\mu\uparrow\mu\downarrow & \end{array}$$

The upshot of all this is that Clark has great freedom of choice in selecting underlying representations for a given tone, a freedom

which she makes full use of. And yet it is apparently never the case that a language realizes all these contrasts even though, according to Clark, they are possible phonological distinctions.

The problem is not unique to Clark, and arises in most versions of level tone theory also. In some ways Clark's five rising and five falling tones are an improvement over the ten predicted in a non-hierarchical five-level-tone theory, but they are offset by the proliferation of level tones, and therefore of total tonal inventory, given below for bi-moraic syllables:

$\uparrow\mu\mu$ $\mu\mu\downarrow$ $\uparrow\mu\mu\downarrow$ $\mu\uparrow\mu$ $\downarrow\mu\uparrow\mu$ $\mu\downarrow\mu$ $\uparrow\mu\downarrow\mu$
 $\downarrow\mu\mu$ $\mu\mu\uparrow$ $\downarrow\mu\mu\uparrow$ $\mu\uparrow\mu\downarrow$ $\downarrow\mu\uparrow\mu\downarrow$ $\mu\downarrow\mu\uparrow$ $\uparrow\mu\downarrow\mu\uparrow$
 $\mu\mu$

Excluding syllables involving sequences of like tone markers (such as $\uparrow\mu\uparrow\mu$) Clark can represent 15 different tones. The Chinese languages, which probably exhibit the largest tonal inventory on mono-syllabic morphemes, never have more than ten (Cantonese) as far as I know.

Since Clark makes use of possibilities of different types within one language (e.g., Chaozhou $\downarrow\mu\uparrow$ and $\downarrow\mu$ are both level tones) it seems unlikely that she can claim that languages draw from this inventory in principled ways (such as using only initial or only final tone markers and so forth).

(d) Floating Tones

Unlike a segmental approach, Clark can easily describe morphemes that consist solely of tone. The problem comes in capturing the

behaviour of such floating tones as those of Cantonese (see section 1.1.1). In a level tone theory Cantonese has a morpheme consisting of high tone only. If added to a syllable that already begins on a high tone the output is a high level tone. If added to a morpheme that begins on a non-high tone, the result is a tone that rises to a high level. This rise results even on mono-moraic syllables, which requires a special realization convention for a rise on a single mora, or positioning of a marker inside the mora.

The following discussion is based on one possible approach to the Cantonese facts suggested by Clark herself (personal communication). The first problem is that Cantonese has four level tones, and Clark can only represent three without running into problems elsewhere; suppose 55 is $\uparrow\mu\mu\downarrow$, 33 is $\mu\mu$ and 22 is $\downarrow\mu\mu\uparrow$; then the remaining 44 must either involve a smaller rise $\uparrow\mu\mu\downarrow$ or perhaps be $\uparrow\mu\mu$. The complete set of underlying representations would then be as follows:

55	$\uparrow\mu\mu\downarrow$	5	$\uparrow\mu\downarrow$
53	$\uparrow\mu\downarrow\mu$	4	$\uparrow\mu$
44	$\uparrow\mu\mu$	3	μ
33	$\mu\mu$		
22	$\downarrow\mu\mu\uparrow$		
24	$\downarrow\mu\uparrow\mu$		
35	$\mu\uparrow\mu\downarrow$		

Now the floating tone is probably made up of the sequence $\uparrow\downarrow$ to trigger the right rise on the preceding syllable and ensure a drop

back into the subsequent syllable. This would also be the residue from deleting the segmental material of a 55 or 35 syllable (see section 1.3.1), another source of these tonal changes.

Consider the consequences of adding this to the 22 tone.

$$\downarrow \mu \mu \uparrow + \uparrow \downarrow$$

The output is a rising tone, which means an inter-moraic \uparrow . There is also an undesirable sequence of $\uparrow \uparrow$ present. Let us therefore write two informal rules:

(1)

$$\uparrow \longrightarrow \emptyset / \text{---} \dots \uparrow$$

and

(2)

$$\uparrow \downarrow \downarrow$$

These two rules will derive the right result for tones 22, 24 and 33:

	22	24	33
	$\downarrow \mu \mu \uparrow + \uparrow \downarrow$	$\downarrow \mu \uparrow \mu + \uparrow \downarrow$	$\mu \mu + \uparrow \downarrow$
(1)	$\downarrow \mu \mu + \uparrow \downarrow$	$\downarrow \mu \mu + \uparrow \downarrow$	$\mu \mu + \uparrow \downarrow$
(2)	$\downarrow \mu \uparrow \mu + \downarrow$	$\downarrow \mu \uparrow \mu + \downarrow$	$\mu \uparrow \mu + \downarrow$

Now consider the result of adding the sequence $\uparrow \downarrow$ to tone 55:

$$\uparrow \mu \mu \downarrow + \uparrow \downarrow$$

There is an unnecessary collection of contradictory markers at the end, so we simplify with rule (3):

(3) $\downarrow \uparrow \rightarrow \emptyset$

This will now account for the remaining tones, sample derivations follow:⁸

	55	53	35
	$\uparrow\mu\mu\downarrow + \uparrow\downarrow$	$\uparrow\mu\downarrow\mu + \uparrow\downarrow$	$\mu\uparrow\mu\downarrow + \uparrow\downarrow$
(2)	—	$\uparrow\mu\downarrow\uparrow\mu + \downarrow$	—
(3)	$\uparrow\mu\mu + \downarrow$	$\uparrow\mu\mu + \downarrow$	$\mu\uparrow\mu + \downarrow$

The point of this illustration is to show that although it is possible to arrive at an analysis of the facts in Clark's framework, in no way can it be called natural. In an autosegmental framework the changed tones in Cantonese are an extremely simple phenomenon requiring no phonological rules at all. In a dynamic tone analysis they appear unintuitive. As far as I can see similar problems will always arise when adding floating tones to any system that uses a mixture of initial, medial and final tone markers; it would probably not constitute a problem in a system like Mende with initial markers only. I should make it clear that I take full responsibility for this analysis; it may well be that more thought would show a more illuminating approach within Clark's theory.

1.7.3 Problems Stemming From Positioning the Tone Markers Among the Segmentals

(a) Deletion of Tone or Segments Only

As Clark points out, her theory shares with autosegmental theory the property of predicting the phenomenon referred to as stability: the invariance of segmentals during tonal changes, and

the invariance of tones during segmental changes. The problem is rather one of formalism. In autosegmental theory one can write rules that refer only to tones since they form a separate tier in which they are adjacent. In dynamic tone theory tones are separated by segmentals which must therefore be allowed for in the rules. For example, suppose a language has a rule which deletes a low tone between two high tones. In autosegmental terms this can be written as:

$$L \longrightarrow \emptyset / H \text{ ____ } H$$

In dynamic tone terms the rule will look something like this:

$$\downarrow \longrightarrow \emptyset / \uparrow_{\mu} \text{ ____ } \mu \uparrow_{\mu_0} \downarrow$$

The need to make reference to the segmentals also makes claims about the sorts of rules one finds in language. For example, Clark gives rules which make reference to syllable structure by permitting intervening vowels or vowel-nasal sequences, but not consonants. Such rules are given on p. 175:

3.2-1 \downarrow -Metathesis (Igbo)

[\downarrow [+syll] ... \downarrow]
NP

s.d. 1 2 3 4, where 3 contains no tone markers

s.c. Transpose 1 and 2

3.2-2 Syllabic Nasal Raising (Igbo)

$$\# \# [+Tone] \begin{bmatrix} + \text{syll} \\ + \text{nas} \end{bmatrix} [+ \text{nas}]$$

s.d. 1 2 3 4

s.c. Transpose 2 and 3

What is at issue is not whether or not such rules exist -- apparently they do -- but rather how they should be evaluated. In dynamic tone theory 3.2-1 is just as simple as rules moving tone markers over any syllable as far as I can see. Yet rules of the latter kind are very common, and rules of the first kind are highly unusual. In autosegmental theory this difference is captured by making rules which act on one tier but make reference to another tier much more expensive than rules operating entirely within one tier, but no such simple criterion is available here.

The fact that allowance has to be made for segmentals in rules affecting only tones has another effect on the forms of rules; in almost every rule variables have to be used and the variables must normally be restricted to [-tone]. For example, Mandarin Like-Tone Marker deletion (Clark's 5-6) has the following form:

$$\begin{bmatrix} +\text{tone} \\ \langle +\text{fall} \rangle \end{bmatrix} \quad \dots \quad \begin{bmatrix} +\text{tone} \\ \langle +\text{fall} \rangle \end{bmatrix} \quad \dots \quad \$$$

s.d. 1 2 3 4 5

s.c. Delete 1

Condition: 2 contains no tone markers

In order to capture the fact that rules affecting [+tone] require

the variables to range only over [-tone] and vice-versa some notion of projection needs to be added to the theory. This would allow reference to tone markers as adjacent in the projection even when they are separated by non-tonal material in the string. This power is also needed for rules that delete (or copy) all the tones or all the segments of a morpheme, as in Mandarin neutral tone (1.4.1) and reduplication (1.1.2). However, there would still be rules that could not be stated on the projection, since their effect is to move tone markers over segmentals, and no effect would therefore be observed on the projection level alone. One rule of this kind is ↓-Displacement in Kikuyu (Clark's 3.3-2, p. 179):

↓ ↓ ↑

1 2 3 4 5

3 replaces 5

Condition: 4 contains no tone markers

Clark claims that movement rules will always be so constrained as not to move one tone-marker over another, but notice that this does not fall out from dynamic tone theory directly: an additional statement to this effect is required. In autosegmental theory, on the other hand, the movement of tones does not take place; movement rules require the power of transformational rules, a power that is essential to Clark, and which she uses liberally. But in autosegmental theory it may be possible to exclude rules of this form, and it is certainly true that in general they are not needed. The fact that tones do not cross over other tones falls out directly

from the WFC which states that association lines may not cross, and without transformational rules there is no other way to effect a cross-over.

(b) Contour Tones on Short Vowels

The existence of contour tones on short vowels constitutes one of the strongest arguments in favour of an autosegmental theory. Clark expresses some doubt as to whether such tones can be contrastive, but it is quite clear that they can in languages like Cantonese. And Clark admits (p. 197) in connection with some data from Haraguchi (1975)

"There is no way to account for the data...within the dynamic-tone theory which I have proposed here, for my assumptions (i) that the mora is the smallest possible tone-bearing unit, (ii) that a tone-marker may not appear inside a tone-bearing unit, and (iii) that two tone markers may not appear at the same boundary prevent me from giving distinct tone representations to the two items kasiwa/mo\ti and harisi\goto."

In cases of clear word-final contour tones on short syllables, such as the difference between Mende ↓kpa — and ↓mba↑ —, low level and rising, the glide is analyzed as the realization of a word-final tone marker that has no other possible way of surfacing except as a glide on the preceding syllable. The claim is therefore that the existence of gliding tones on word-final syllables can be explained by means of plausible phonetic realization rules for the tone markers. Word-internal glides like those in the Japanese example given earlier are much harder to account for since no obvious realization rule will account for them.

Clark tentatively suggests that the theory could be extended for

such cases to allow tone markers between initial and rhyme, but this would greatly increase the inventory of possible tones and therefore considerably weaken the theory. No such undesirable consequences attach to describing contour tones on short syllables in an autosegmental framework.

Before closing this discussion of dynamic-tone theory, let me briefly show how the apparent elegance of one of Clark's analyses fails under closer scrutiny. In her analysis of Mandarin on p. 137 she gives a rule which accounts for the change of third tone from 214 before a pause to 21 elsewhere. The rule applies to the underlying representation $\downarrow \mu \mu \uparrow \mu$ and is formulated as follows:

$$\begin{array}{rcccccc}
 & & [& \mu & \mu & \mu &] & \dots & \$ \\
 \text{s.d.} & & 1 & 2 & 3 & & 4 & & 5 \\
 \text{s.c.} & & & & \text{Delete 3} & & & & \\
 & & & & \text{Condition: } 4 \neq \emptyset & & & &
 \end{array}$$

It therefore converts $\downarrow \mu \mu \uparrow \mu$ to $\downarrow \mu \mu \uparrow$, a low level tone.

Unfortunately this rule will also mistakenly delete final nasals in words like zhōng, producing:

$$\downarrow \text{zhoo} \uparrow \text{ng} \longrightarrow \downarrow \text{zhoo} \uparrow$$

If instead the second mora is deleted, we get the right segmentals, but the wrong tone:

$$\downarrow \text{zhoo} \uparrow \text{ng} \longrightarrow \downarrow \text{zho} \uparrow \text{ng}$$

Clark has tried to achieve a tonal effect by means of a segmental

deletion, but the simplicity of the rule is valueless because it fails to achieve even descriptive adequacy. In a recent unpublished paper, Clark (1978b) has revised her analysis of Mandarin so that this problem is avoided, but I have included it here to make the point that complications of this kind simply cannot arise in autosegmental theory, which properly separates tone and phonemes.

To sum up, there are a number of serious problems with Clark's dynamic-tone theory, and it offers few advantages over autosegmental approaches. Unless far-reaching changes that vastly increase the power of the theory are made it fails to describe some phenomena, such as contour tones on single moras, and predicts phenomena that do not occur, such as tonal conditioning by the number of mora in the following syllable. Such a detailed exploration of the consequences of a dynamic-tone theory is well-worth undertaking, but we must unfortunately conclude that Clark has failed to prove the pre-eminence of dynamic-tone theory over autosegmental theory.

NOTES TO CHAPTER ONE

¹ Paul Kiparsky has pointed out to me that in fact a lexical rule changing the second of two tones to a high 5 would achieve the right result. This would fail in the case of stopped syllables, which only have one tone underlyingly and therefore require the addition of a high 5 tone.

² Some borrowings, such as mǐ 米 'metre' do not reduplicate unless fully naturalized (e.g., bàng bàng 磅 'every pound').

³ In some dialects at least the tone never changes in bisyllabic forms. In such dialects the reduplication rule must be conditioned by the number of syllables as to whether or not the tone is also reduplicated.

整	，	齐	，	zhěng zhěng qí qí	'neat'
扁	，	原	，	biǎn biǎn yuán yuán	'plain, flat'

⁴ Neutral tone is shown by the absence of a tone mark.

⁵ There are also [11] tones which are not neutral tones, for example the sandhi form of /55/, see 3(b). These differ from neutral tone in bearing stress.

⁶ This is somewhat misleading because there is an additional process that eventually lowers the pitch of the third and subsequent syllables (see section 5.2). However, the argument still goes through.

⁷ Typo in Osburne (1979) says [- sonorant]. Osburne's thesis has [+ sonorant].

⁸ Since finishing this chapter I have realized that I have omitted the most glaring problem with a dynamic tone analysis of Cantonese changed tone: the rising tone on single mora syllables. In order to change μ to a rise it will be necessary to add a mora:

$$\mu + \uparrow\downarrow \longrightarrow \mu\mu + \uparrow\downarrow \xrightarrow{(2)} \mu\uparrow\mu\downarrow$$

There must therefore be a rule of the following form:

$$\emptyset \longrightarrow \mu / \$ \mu \dots \uparrow\downarrow$$

This further obscures the unity and simplicity of the phenomenon that an autosegmental analysis expresses so elegantly.

2. THE INTERACTION OF TONE WITH OTHER PROSODIC SYSTEMS

There has been much discussion of the lack of interaction between the tonal and segmental levels, and indeed their relative independence is one of the major arguments for regarding tone as suprasegmental. There has been relatively little discussion of the other side of the coin: the frequency with which lexical tone interacts with other prosodic systems such as stress, and the importance of tone in metrical verse. In a theory in which tone and stress, or either one of the two, are features on segments there is no way of explaining the fact that tone and stress interact more commonly than tone and vowel backness, or that verse may exhibit tonal patterns but rarely patterns of vowel height. On the other hand if both tone and stress are suprasegmental their indifference to the segmental level is explained, and their sensitivity to each other may also be explained once the relationship between different suprasegmentals is understood. In this chapter we will examine the interaction of tone and stress in three Chinese languages, and the tonal patterns of a traditional verse form. We will discover that if tones themselves are assigned a branching metrical structure their stress conditioning can be subsumed under general metrical conventions proposed for other languages such as English.

2.1 Tone and Stress

2.1.1 Mandarin

Mandarin has final stress in both words and phrases made up of full-toned morphemes.¹ The secondary stress falls on the initial

syllable, with intervening syllables carrying lower levels of stress.

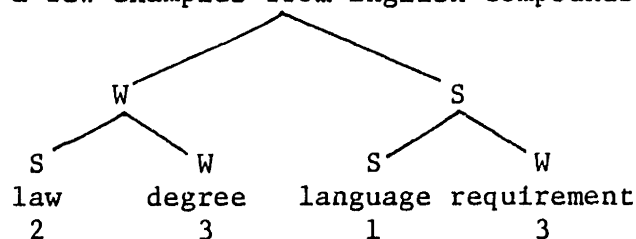
2 1 gāoxìng	'happy'
2 3 1 túshūguǎn	'library'
2 3 3 1 gòngchǎndǎngyuán	'Communist party member'

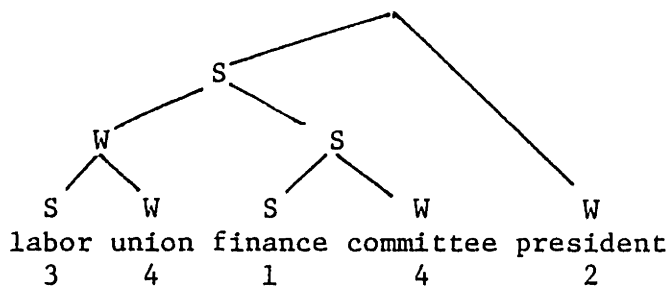
In the metrical theory developed in Liberman & Prince (1977) stress is represented by means of binary branching metrical trees, the nodes of which are labelled strong (S) and weak (W) such that sister nodes always carry opposite labels. The main stress falls on the syllable dominated by an unbroken branch of S nodes, and lower levels of stress can be read off by using the principle that every S-W pair of sister nodes defines the relative prominence of the sub-units of the tree dominated by those nodes. (For convenience the following algorithm can be used:

The stress number of a terminal node is equal to the number of nodes that dominate the lowest W dominating that node, plus one. (Adapted from L&P: 258)

But it should be remembered that metrical theory makes the claim that stress is a matter of relative prominence rather than numerical values, and the algorithm is intended only to allow comparison with standard notation.)

Let us give a few examples from English compounds:





The labels S-W are assigned by means of the Word Rule (L&P: 268).

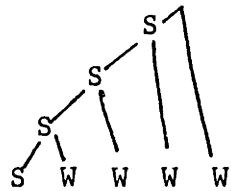
L_1 : In a pair of sister nodes N_1N_2 , N_2 is S iff it branches. The geometry of the tree in English compounds is read directly off the syntax.

Now consider the sort of stress patterns that can be described by such a metrical structure. Suppose that we ignore syntactic structure and look at the trees as abstract entities. The simplest kind of tree would be unidirectional in its branching, giving two possibilities, and could be labelled by the rule L_1 given above, or perhaps by the formally identical but mirror-image rule:

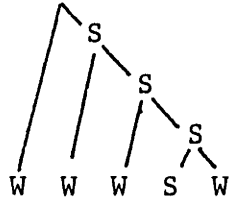
L_2 : In a pair of sister nodes N_1N_2 , N_1 is labelled S iff it branches.

This produces the following four possibilities:

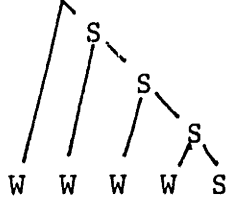
Initial stress: Left-branching, L_1



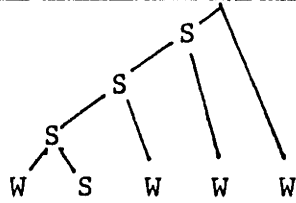
Penultimate stress: Right-branching, L_1



Final Stress: Right-branching, L_2

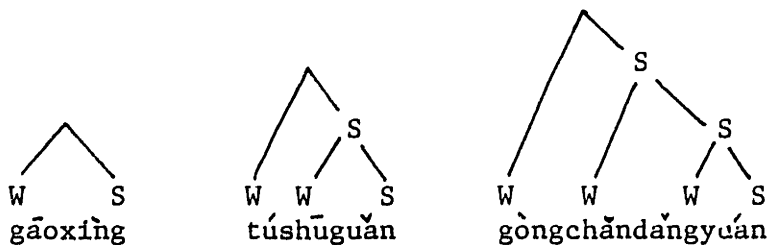


Second syllable stress: Left-branching, L_2



All of these are found in the languages of the world. For example, word-initial stress is found in Hungarian, Latvian, Papago and Chitimacha; penultimate stress is found in Polish, Quechua, Swahili and most Aztec dialects; final stress is found in French, Farsi and Mandarin, and second-syllable stress is found in Lakota and Araucanian. Furthermore, the position of secondary stress is correctly predicted by the above structures to be at the opposite end of the word from main stress. So in Polish which has penultimate main stress the secondary stress is initial; and in Mandarin, with final stress, secondary stress is also initial.

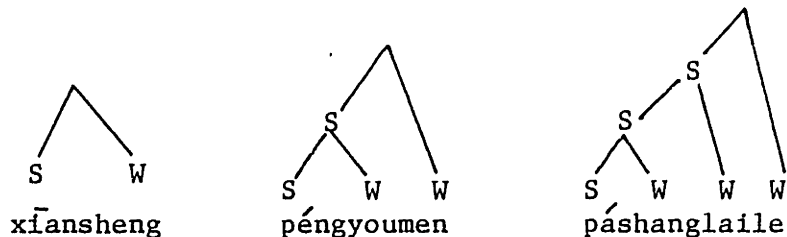
Having briefly given some background to the theory being assumed here, let us return to Mandarin, which has final main stress and initial secondary stress. This is one of the four stress patterns produced by unidirectional trees labelled by L_1 or L_2 , so we can say that stress in Mandarin is assigned by right branching trees labelled by convention L_2 , and this will produce:



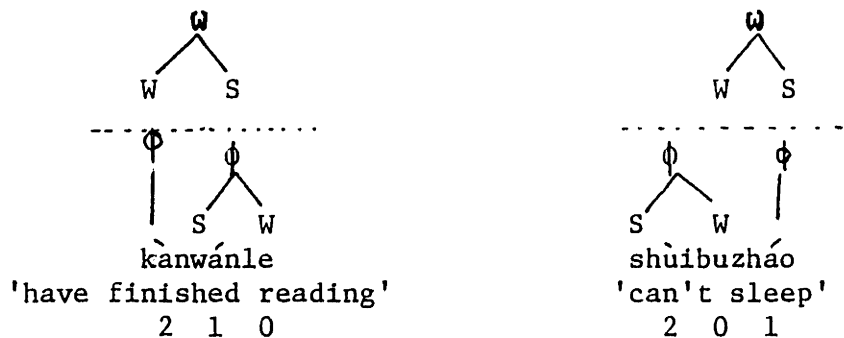
However, there are also many words and phrases with initial stress:

1	0		
xī	sheng	'teacher, mister'	
1	0	0	
péng	you	men	
		'friends'	
1	0	0	0
pá	shang	laile	
			'jumped up'

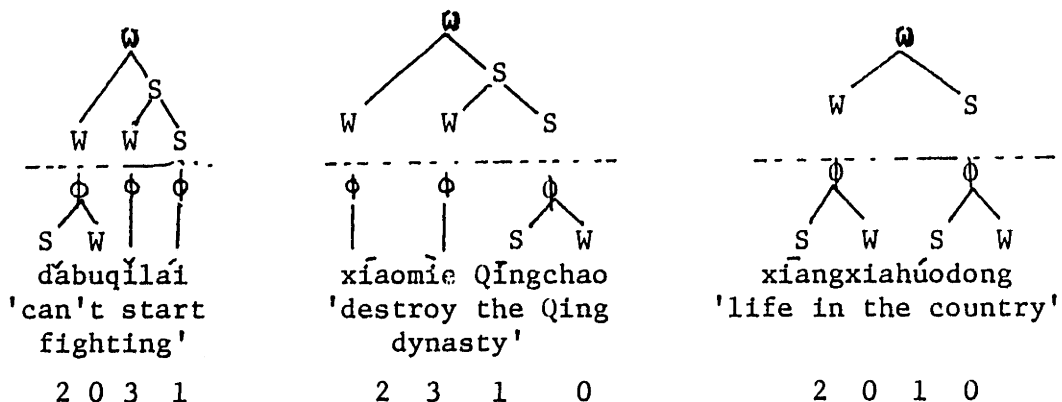
These differ from the earlier examples not only in the position of main stress but also in the amount of stress on the other syllables: in all these cases the non-initial syllables are completely unstressed and surface in the so-called neutral tone, which we have argued in section 1.2.1 is the manifestation of a toneless syllable.² We will maintain that each of the above words consists of a single metrical 'foot,' and that feet may themselves be grouped together into words. The word tree then assigns stress to the foot, rather than to the terminal nodes of the feet. By convention the stress of a foot is then realized on the Designated Terminal Element of that foot, i.e., the node dominated by S nodes only. This means that a foot will have a single stressed element, and the rest will have zero stress. Where the word consists of a single foot, as in péngyoumen, the DTE of the foot will have the main word stress. In Mandarin the DTE of a foot is always the first syllable; it must therefore be the case that feet are left-branching and labelled by convention L_1 . This will produce the following feet:



In this instance, since each of these are words, there is a degenerate word tree that does not branch. However, when a word is made up of several feet the interaction is more complex. The feet themselves may or may not be branching. For example, when there are no neutral tones in the word then each syllable constitutes a foot in itself, and the word tree assigns the stress pattern [2 (3...3) 1]. But there may be neutral toned syllables involved, in which case the stress pattern is determined in two stages. First feet are constructed, then a word tree is built connecting the feet, and assigning main stress to the last foot. Consider the following words:



Notice that the feet are left branching and labelled SW, while the word tree is right branching and labelled WS. This is more apparent in longer words:



If the lexical entry for 'know' has no tone on the second syllable this is very hard to explain. A similar argument can be made from the existence of optional doublets, one with neutral tone and one with full tone:

xī wáng ~ xī wang 'hope'

These cases suggest that we need some sort of diacritic to mark those syllables which show up with neutral tone, rather than an absence of tone being sufficient. A stronger argument comes from the existence of words involving third tone morphemes.

Consider a word like:

láo hu 'tiger' or xiáo jie 'miss'

In both these cases the basic morphemes are all third tone, and presumably native speakers analyze these forms into morphemes, since they all occur commonly either freely (lǎo 'old'; xiǎo 'small') or in other forms (dà jiě 'oldest sister'; hǔ pí 'tiger's skin, bold exterior'). However, the surface forms show second tone followed by neutral tone, suggesting that the third tone sandhi has applied. This means that the final syllable must have third tone to provide the context for sandhi. Since the occurrence of neutral tone is unpredictable, no phonological rule can delete the tone, and we need to assume a diacritic marker, say [-stress], that in turn will trigger tone deletion. Next we consider a different set of cases.

Given an autosegmental analysis of tone we would expect that a language could have both toneless and segmentless morphemes, and

indeed examples of both kinds exist in Mandarin (see 1.1.2 and 1.2.1). What concerns us here are toneless morphemes, which are of two main types: sentence particles, and some suffixes such as locatives and aspectual markers. These also show up with neutral tone, and must be linked to a preceding tone-bearing syllable in a foot. If these are toneless (as is not only sensible but necessary in some cases (see 1.1.2)), rather than marked redundantly [-stress], then the conditioning for foot formation is apparently lack of tone, and we can see that tone deletion (as in láo hu) must precede foot formation (but follow third tone sandhi).

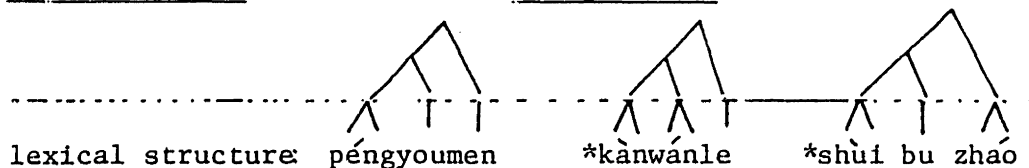
We now have two sources for neutral tone, and there is yet another case. There are words, like pronouns, that undoubtedly have associated tone in the lexical entry, but that can cliticize onto the preceding word in certain syntactic environments. After cliticization they show up with neutral tone (e.g., Tā kàn ta yì yán 'He looked at her for a moment'). No lexical information will deal with this case because it is syntactically conditioned. The most obvious way of capturing the facts would be to propose a syntactically conditioned rule of foot formation that mimics cliticization by attaching the pronoun leftwards to the verb, and then assume that foot internally tone is deleted on all but the designated terminal element of the foot. But now we have two different conditioning factors for tone deletion: [-stress], and foot internally. Is there a way of collapsing these? The answer of course is 'Yes': [-stress] conditions not only tone deletion, but also foot formation. Suppose that we skip this intermediary diacritic, and enter a partial foot

structure directly in the lexicon. Neutral toned (and toneless) syllables will be distinguished from fully toned syllables by having non-branching as opposed to branching structures associated with them.

We can then state the conditions on foot formation geometrically. Suppose that feet must have a single direction of branching, and in Mandarin that means left-branching. If syllables already have associated structure, that structure must fit into the foot template. This means that only the first syllable of a foot can be one with lexical branching, because if any of the others branch they will violate the leftwards character of the foot:

Permitted foot:

Illegal feet:

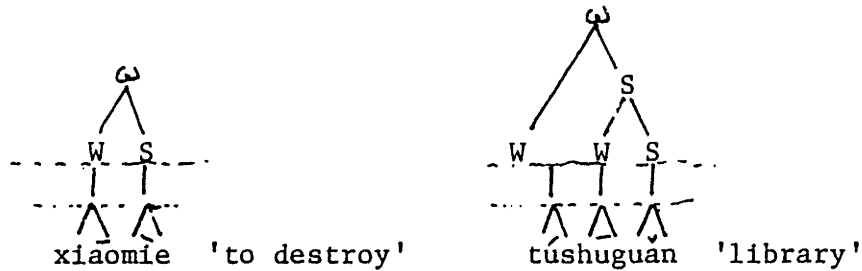


More formally, feet are subject to the condition that the initial node may branch, but no other node may branch.

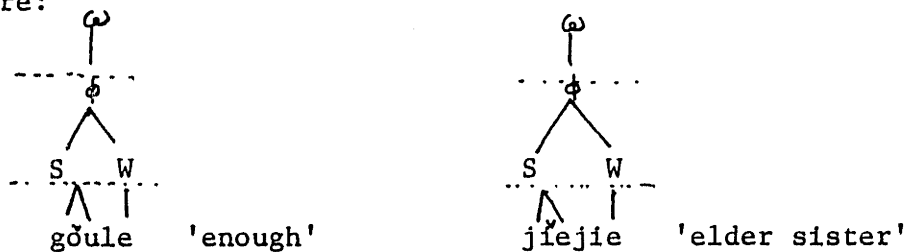
Tone deletion can now be stated as a foot-internal process, affecting all but the DTE of the foot, and leaving only one tone per foot. This is apparently a common restraint in languages: Shanghai has a similar restriction to one tone per foot, and so does Capanahua, a S.American Indian language (Safir 1979). Destressing in English is frequently a result of de-footing, and the defooted vowels are then reduced to the neutral vowel [ə], as if only one vowel per foot could have marked values for the vowel features. The equivalent of English destressing in Mandarin is the process of cliticization that affects

object pronouns, defooting them and attaching them in a left branching foot. In section 5.1 we will see certain phonological processes in Mandarin whose context turns out, in this framework, to be foot-internal.

If the preceding discussion is right and feet are formed from a single branching node over any number of non-branching nodes into a left-branching structure, then there must be lexical entries of three kinds. The normal case would be for each syllable to have full tone and a branching structure associated with it:



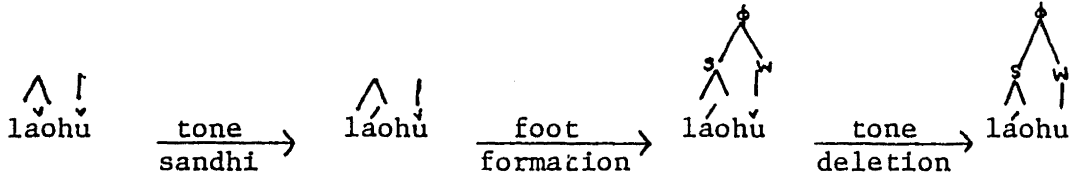
No feet can then be formed, and the word tree will treat each syllable as an independent foot and assign final stress. The next common kind would be words including a syllable with no tone and an unbranching structure:



Each of these will be formed into a single foot by the foot formation rules, and the degenerate word tree will result in main stress falling on the first syllable of the foot.

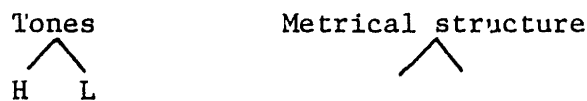
The third kind involves syllables with full tone but unbranching structure; since the structure is unbranching they will be incorporated

into a foot, and the tone will then be deleted foot-internally:



The fact that tone sandhi has applied is evidence that the full tone was present initially, but since it surfaces in neutral tone it must have been unbranching and therefore been absorbed into a foot.

The fourth logical possibility -- no tone but a branching structure -- does not exist. It is clear that the unmarked cases are the first two: [branching structure + full tone] or [unbranching structure + no tone]; another way of viewing the process of tone loss is then to say that at the surface realization of tone requires association with a branching structure, and that if this structure is lacking the tone cannot surface. This is a particularly satisfying way to look at things in the case of the contour tones, since there is then a congruence between the branching metrical structure and the branching of the tonal sequence:

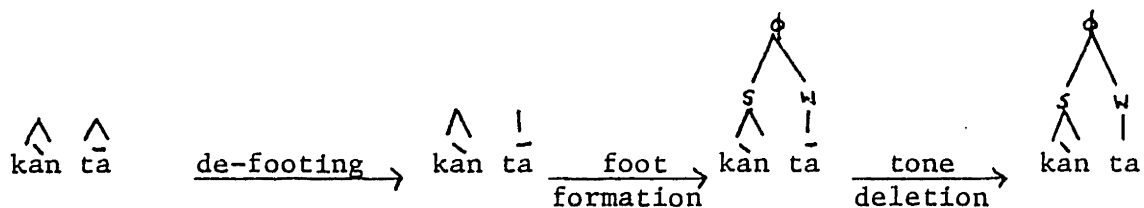


This suggests a stronger hypothesis. Suppose that not only the contour tones but also the level tones are sequences of tonemes, so that the first tone, which is high level, is HH, and the third tone, which is low level, is LL. Then it is possible to say that stress trees are built on the tonal structure, and since this is normally branching tone-bearing syllables will be treated as metrically branching also

unless specially marked to the contrary. Non-toned syllables (with underlying tones) will, as before, be entered with a non-branching structure in the lexicon.

From now on we will assume that stress trees are built on Tones in this way. One result is that stressed syllables must always have two tonemes underlyingly (HH rather than H), and it is interesting to see that this mirrors the intuitions of the traditional sinologists, who always denoted tones with two digits (except on stopped syllables) implying that the starting point and the endpoint of the tone had to be stated independently even when identical: 55, rather than 5.

Lastly, let us give an example of cliticization:

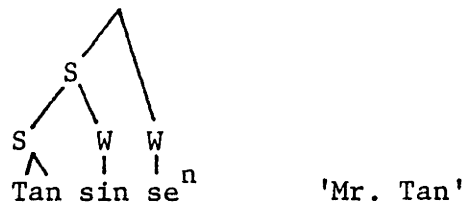


To summarize, morphemes may or may not have tone, and they are also accompanied by partial foot structure in the lexicon: branching or non-branching nodes. Toneless morphemes are always non-branching, but tone-bearing morphemes may be branching (most common) or non-branching (marked). Feet are formed left to right by the usual conventions, labelled SW. The word tree links feet right to left, and is labelled WS. Third tone sandhi precedes de-footing (cliticization) and perhaps all metrical rules. Tones are deleted foot internally, and tone then spreads from the DTE of the foot over the rest of the syllables in that foot.

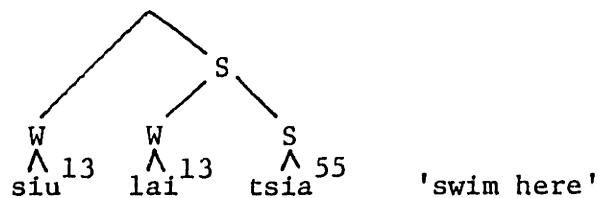
2.1.2 Amoy

Amoy has essentially the same stress patterns as Mandarin, with the exception of an additional special pattern before one particular suffix. Elsewhere, left-branching feet are formed over non-branching syllables, and labelled SW by L_1 ; these are linked into a word tree that is right-branching and labelled WS by L_2 .

Foot:



Word:

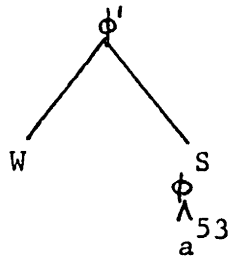


Bisyllabic words that are single feet therefore have the stress pattern 1-0, while those that are two feet have the pattern 2-1.

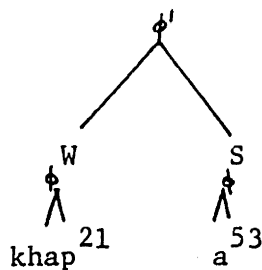
The additional twist comes when we examine the suffix a⁵³, which has, among other things, a diminutive meaning, and has been discussed in connection with its special tonal effects in section 1.4.3.

Bisyllabic words ending in this suffix have the stress pattern 3-1, the only time this pattern occurs in the language. How is this to be explained? I would like to suggest that this suffix is special in requiring that the preceding syllable join it in a special W-S foot as distinct from the W-S pattern formed by a word tree. This can be captured by entering the suffix in the lexicon with an associated

foot structure including an empty node to be filled by the preceding syllable:



The difference in stress between the weak member of a usual S-W foot, which has 0 stress, and the weak member of this foot, which has 3 stress, is taken to be a side-effect of the fact that these special feet produce the only W members of a foot which have full tone (and are themselves branching feet), and they are therefore perceived as having some stress; W members of feet with no tone are perceived, on the other hand, as having 0 stress. This special foot is in fact made up of two normal feet, since both syllables have full tones and branching structures. The affix apparently triggers a kind of Chomsky adjunction of the preceding syllable as weak sister under a higher 'super-foot' node:

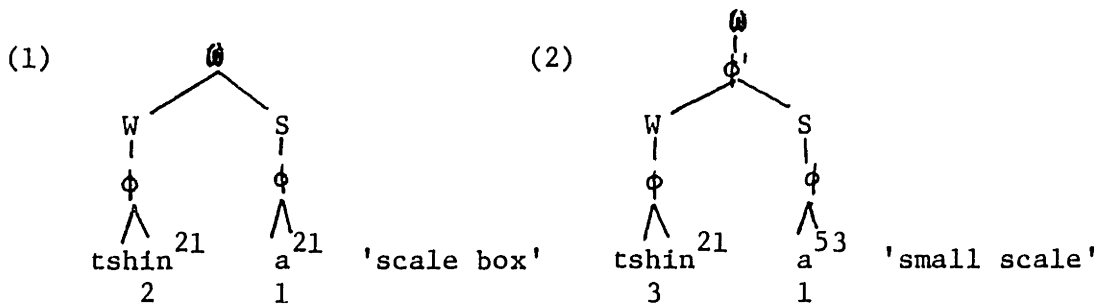


Complex feet of a related kind have been argued for in Passamaquoddy and Seneca by Stowell (1979: 56ff).

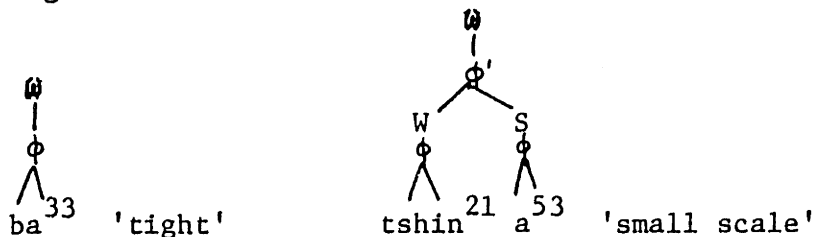
We mentioned earlier that this suffix is special in several ways:

- (i) stress pattern 3-1
- (ii) treated metrically as a single syllable
- (iii) triggers deletion of melody on preceding syllable
- (iv) triggers re-syllabification and stop voicing

Given the approach outlined above it is now possible to give a unified explanation of these various phenomena. The stress pattern 3-1 is the result of the introduction of an extra level of embedding in the metrical tree. Using ω to represent the word level, compare (1) a normal word made up of two fully-toned morphemes with (2) one including the affix a⁵³:



Using the algorithm in 2.1.1 and counting up from the W dominating tshin²¹ in each case, we derive the right values of the non-primary stress. At the same time we explain why words like [tshin⁵⁵ a⁵³] 'small scale' are treated metrically as a single syllable: at the level of immediate constituents they, like monosyllabic words, consist of a single foot:



The following is a fragment of a children's chant from Taiwan (James

Huang, personal communication):

1	2	3	4	5	6	7	
ginjio	lingling	jia	linging				
banana cold-cold, eat longan							
1	2	3	4	5	6	7	
linging	bo	ba	jia	yu-a	ba		
longan no meat, eat grapefruit meat							

The first line is typical of the chant and shows the seven-syllable pattern. The second line has eight syllables, but one of them is the a⁵³ suffix which forms part of the word yu-a 'grapefruit': these two syllables are counted as a single syllable and the line therefore fits the pattern. There is thus clear evidence for the metrical identity of a normal syllable, and a normal syllable plus a⁵³. If each of these constitutes a single foot the equivalence is explained.

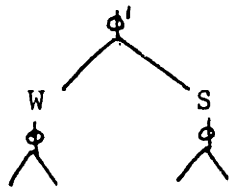
The melody deletion caused by this suffix can now also be understood: foot-internally only one melody is permitted, and that is the melody of the S terminal node (the DTE). Just as in a normal foot the melody on the second, W, node is deleted, in this special super-foot the melody on the first, W, node is deleted.

Lastly, the processes of resyllabification and stop voicing can be formulated as foot-internal processes. There are contrasts of the following kind between ordinary words and words involving a⁵³:

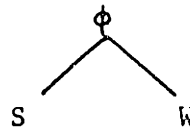
	tshin ²¹	a ²¹	→	tshin ⁵³	a ²¹	'scale box'				
<u>but</u>	tshin ²¹	a ⁵³	→	tshin ⁵⁵	na ⁵³	'small scale'				
	hang ⁵⁵	khap ²¹	a ²¹	a ⁵³	→	hang ³³	khap ⁵	a ³³	a ⁵³	'clam box'
<u>but</u>	hang ⁵⁵	khap ²¹	a ⁵³	→	hang ³³	khab ⁵	ba ⁵³	'clam'		

Before a⁵³ there is a resyllabification or gemination of the final consonant of the preceding syllable; at the same time voiceless obstruents become voiced. This voicing not only happens before a⁵³ but also optionally before V-initial neutral toned syllables, and in Mandarin at the beginning of neutral toned syllables, so that the unaspirated voiceless [t] denoted by orthographic [d] of zhīdao 'know' is actually a voiced stop when dao is in neutral tone. Let us look at when (i) resyllabification and (ii) unaspirated stop voicing occur in Mandarin and Amoy. In Amoy we have two kinds of feet to consider, but in Mandarin only one:

Before a⁵³ (Amoy only)



Before neutral tone



		resyllabn.	voicing	resyllabn.	voicing
<u>Amoy:</u>	C + V	+	+	0	0
	V + C		DNE	—	—
<u>Mandarin:</u>	C + V		DNE	0	N/A
	V + C		DNE	—	+

0 = optional, N/A = no such forms exist

To give examples of some of the cases in the right-hand column, starting from the top: in Amoy the suffix e 'of' triggers both resyllabification and voicing:

$$\text{tek}^{21} \text{ e} \longrightarrow \text{teg}^{21} \text{ ge}$$

In some other cases it is apparently optional, such as with the classifier e¹³:

$$\begin{array}{l} \text{be}^{53} \text{ tsit}^{54} \text{ e}^{13} \longrightarrow \text{be}^{53} \text{ tsit e} \\ \text{or} \longrightarrow \text{be}^{53} \text{ tsil-le}^3 \end{array}$$

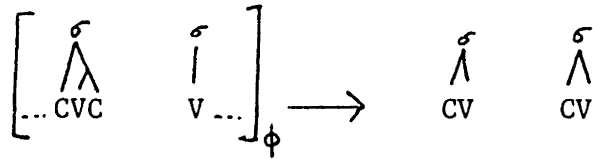
(The l denotes a kind of flap.)

In Mandarin vowel-initial particles also have this effect, although since no Mandarin words end in voiceless stops only the resyllabification effect is apparent (Chao, 1968: 803):

$$\text{rén a} \longrightarrow \text{rén na}$$

On the other hand the voicing effect shows up when the C begins the second syllable, as in zhīdao; whether there is accompanying resyllabification is unclear.

It is clear from the above table that resyllabification applies (obligatorily or optionally) whenever a consonant-final syllable precedes a vowel-initial syllable in the same foot, whether that foot is SW or WS. Resyllabification produces a structure that conforms more closely to the preferred syllable structure for all Chinese languages, CV:



For a more formal statement of such a rule in English see Gray (1979: 178-81). This rule must be stated to apply foot internally, since it fails in words like tshin⁵³ a²¹ 'scale box' that are composed of two stress feet.

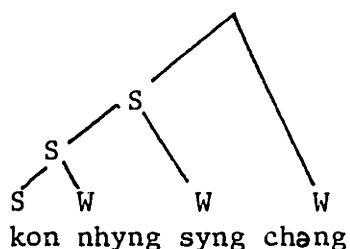
In section 5.1 and 5.3 I will show that once the rule of resyllabification is stated in this way for both Mandarin and Amoy, the voicing facts can be very simply accounted for by different laxing rules in the two languages. (That the rules must differ can be seen from the fact that only Mandarin voices the initial stops of neutral toned syllables.)

Notice that only the notion of foot allows an explanation of these phenomena; no appeal to surface stress patterns would succeed, since the resyllabification takes place in words with the quite different 1-0 and 3-1 patterns, but not in words with 2-1 stress. Furthermore the notion of foot immediately explains the special treatment of the suffix a⁵³ in verse, and the fact that melody deletion takes place in both 3-1 and 1-0 stress patterns, but not 2-1. For our last look at stress in a Chinese language we will briefly summarize the superficially quite different but very simple facts of Shanghai.

2.1.3 Shanghai

We have already noted, in 1.4.2, that Shanghai has a process

that deletes the tones of all but the initial syllable of a word. The tone of this syllable then spreads over the remaining syllables. The preceding discussion of stress in Mandarin and Amoy should immediately suggest a way of explaining this tone deletion: we have seen that it is commonly the case that foot-internally only one tone, that of the DTE, is permitted, and all others are deleted. If we therefore treat Shanghai words as consisting of a single left-branching foot labelled SW, the first syllable will be the DTE and the tone on all others will be deleted:



This means that the morpheme structure conditions of Shanghai require that all words begin with a branching syllable and are followed by unbranching syllables; the foot formation rules build a single foot, and the tones are then deleted.

The obvious question is whether the facts of Shanghai stress bear out this proposal, and unfortunately we cannot be sure. Sherard (1972) discusses only contrastive stress, and none of his examples provide any clues as to where this would fall in a polysyllabic word. Non-contrastive word stress is not mentioned, which suggests that it is not prominent in its phonetic effects. I have asked Sherard whether words have initial stress, as predicted by the above approach, and he was uncertain, so final resolution of this question must await

further data. Certainly nothing I have been able to find in the literature makes the above proposal implausible, and there is some highly suggestive supporting evidence in an article by Kennedy (1953) in which he discusses Tangsic, a Wu dialect closely related to Shanghai. Tangsic shows an apparently identical phenomenon of tone-spreading from the first syllable of a word (but not a phrase), and Kennedy says (p. 372):

"It is phonetically demonstrable that in (pattern) A (i.e. the word pattern) the louder stress is on the first syllable, in (pattern) B (i.e. the phrase pattern) on the second."

Before leaving the question of stress I would like to add one last example of the role played by the foot in the realization of tone. I have shown that the foot can be considered the domain of tone in three Chinese languages: in each case feet are normally constructed over one tone-bearing (i.e., branching) syllable and any number of adjacent toneless (non-branching) syllables. When under certain circumstances feet are built over two or more tone-bearing syllables the tone of all but the first is ultimately deleted, so that on the surface the tone-foot mapping is always one-to-one. I will now give some evidence from a totally unrelated language that reinforces the supposition that the foot is the domain of tone. The evidence comes from the South American Indian language, Capanhua, and the data and arguments are taken from Safir (1979: 102).

The facts are as follows:

1. Stress falls on the second syllable if closed, otherwise on the first.

2. H tone is assigned to the second syllable if closed, otherwise to the first two syllables.

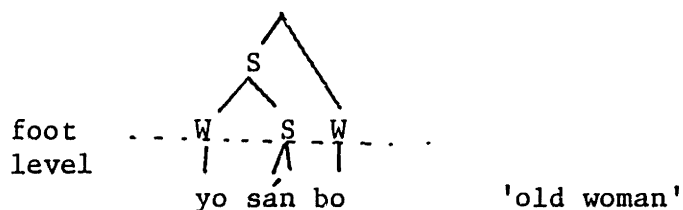
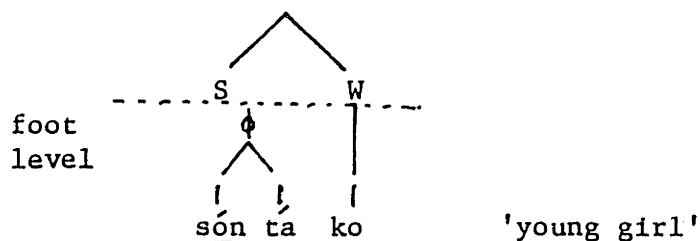
čóskó 'four' where = stress / = H tone
 hisís 'ant'

The rules for building feet are as follows:

1. Only second syllable branching rimes are projected as branching.
2. Form feet SW from adjacent light rimes, from left to right.
3. N_1 S iff branches (word tree).

Then Metrical Tone Rule I

Assign H tone to the S foot.



Again then Tone is shared throughout the foot, showing up on both syllables of a branching foot.

2.2 Traditional Chinese Verse⁴

2.2.1 A Metrical Analysis

During the T'ang period (618-906 AD) Chinese poets perfected a very rigid verse form called lǜ shī 律詩 or regulated verse. In addition to constraints on the numbers of lines and the way they

interrhymed there were very strong restrictions that applied within the line. There were two closely related forms of this verse; in one each line had exactly five syllables, and in the other each line had exactly seven syllables. The interesting part of the principles of versification in lǜ shī is that there were also a set of principles governing which tone could occur at various positions in the line. The four tones of Middle Chinese were divided into two kinds; one tone was the píng 平 or 'even' tone, while the other three were zè, 仄 or 'oblique'. The verse rules specified where a level tone was permitted and where an oblique tone was permitted, and since the level tone was by far the most frequent in the lexicon the level/oblique distinction split the vocabulary roughly 50:50.

The bulk of this section will be concerned with formulating the rules governing the tone distribution, and relating these to other facts about the way verse is recited, and the way the syntax is related to the metrics. Before we get involved in this quite intricate problem a brief excursus on the likely phonological forms of level and oblique seems in order (see also 3.2). Middle Chinese had four tonal categories called by their traditional names píng, shàng, qù and rù. The last, rù, occurred only with syllables ending in a stop consonant, while the other three occurred only with sonorant-final syllables. Attempts have been made to reconstruct the values of these tones based on contemporary descriptions (so far attempts to work backwards from the modern reflexes of these tones have foundered on the rock of the extreme diversity of those reflexes; for example, the 18 dialects taken as representative of all major dialect groups in HYFYCH show

the following reflexes of the ping tone with (A) voiceless and (B) voiced initial consonants:

(A)	55	} level	(B)	55	level
	44			35	} rise
	33			24	
	53	} falling	12		
	42			52	} fall
	31			42	
	21			31	
	213	} fall-rise	443		
	212			21	

This includes level, rising and falling tones at every conceivable level.) The two main attempts at historical reconstructions are Mei (1970) and Pulleyblank (1978). Both are in agreement on two points:

- (i) Ping was used to transcribe Sanskrit long vowels
- (ii) Shang was used to transcribe short vowels

However, they disagree on whether Qu was long or short, whether Shang was high level or rising, and whether the level/oblique distinction was one of low/high or long/short. If qu was long, then obviously the distinction was not level:long versus oblique:short. Mei notes that qu was sometimes used to transcribe long vowels, and concludes that it must therefore have been longish. He further analyzes the shang tone as high, and is therefore led to characterize the level/oblique distinction as one of low/high.

Pulleyblank on the other hand notes that qu was sometimes used at an early stage to transcribe short vowels; he argues that qu had

a final voiced laryngeal h, while shàng had a final glottal stop. Since rù clearly ended in a stop consonant he then points out that píng was the only category ending in a clear vowel, and as such might be expected to be longer than the Oblique categories. For Pulleyblank, then, the distinction is long/short.

Whatever the actual values of level/oblique may have been turns out not to be relevant to our present concerns, and is introduced here as background only. The reason it does not matter is that we will see shortly that for any given abstract metrical pattern there were two exactly opposite realizations; for every line that began with a level tone followed by some pattern of level and oblique there is another possible line that began with an oblique followed by the same pattern of oblique and level. The only asymmetry lies in a strong preference in the so-called recent style (jinti 近體) for the rhyming syllables to be in level tones, which means that even-numbered lines usually follow patterns that end with a level tone.

Returning then to the main topic of this section, the patterns of regulated verse are interesting because they show that just as in the more familiar verse of the West stress interacts with the metre, in the verse of China tone interacts with metre. If tone is a suprasegmental phenomenon this is unsurprising, but if it were a segmental feature there is no reason why it should interact with metre any more often than vowel height or place of articulation. We will show that a single metrical structure governs five things:

- (i) the tonal patterns
- (ii) permitted deviations from the canonical tonal patterns

- (iii) recitation rhythm
- (iv) syntax: metre mapping
- (v) rarity of neutral tones

These five phenomena were normally discussed separately by students of lǚ shī, but we will show that it is not coincidence that a particular recitation rhythm is associated with the verse lines, or that variation is permitted in certain positions but not in others. On the contrary all these fall out automatically given a notion of metrical structure.

There are four possible patterns for each length of line, and these reduce formally to two possible patterns, each of which may begin with either a level or an oblique tone. The seven-syllabled patterns can be viewed as being composed of the five-syllabled patterns with an additional pair at the beginning of each line, so that the total can be reduced to pattern A and pattern B, each of which has a longer and a shorter version each of which can occur with initial level or oblique:

- | | | |
|-----|---------------|---------------|
| (1) | A | B |
| | - - - v v | - - v v - |
| | v v v - - | v v - - v |
| | v v - - - v v | v v - - v v - |
| | - - v v v - - | - - v v - - v |

(- = even-toned syllable, v = oblique-toned syllable.)

The extent to which this ideal pattern was actually followed is unclear and will be discussed in some detail later; for the moment we will simply point out that the first and, to a lesser extent, the third positions in a five-syllabled line were relatively 'free'

(and the first, third, and sometimes fifth in a seven-syllabled line).

When recited aloud the verse was read in a clear iambic rhythm, and there were three different possibilities for reciting the last three syllables:

(2)

	<u>Five Syllables</u>	<u>Seven Syllables</u>
(a)		
(b)		
(c)		

How are we to begin to explain these two sets of facts? It is instructive to take a brief look at a traditional approach to the problem, which was to work out what we might call tonal templates: schemata for the tonal alternations which fully or partly determine the permissible patterns, Downer and Graham (1963) proposed a famous schema of this type in which the possible patterns for a seven-syllable quatrain were given as follows:

(3)

	1	2	3	4	5	6	7
		A		B	-	A	v
		B		A	v	B	-
		B		A	-	B	v
		A		B	v	A	-

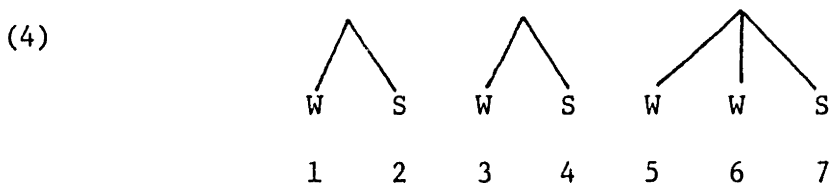
(Positions 1 and 3 are unspecified because they are 'free'.)

The assumption was that this pattern was learned by the poet and poetry reader, and learned quite separately from the other constraints, such as the recitation rhythm. The freedom of the first and third positions

was stated as a fact and no explanation was attempted.

Recently in a couple of interesting papers Matthew Chen (1979, 1980) has explored the possibility of unifying the various properties of lǜ shī by assigning a metrical structure to the line. Below we recapitulate (in a somewhat different form) some of the arguments for such a metrical approach, and then continue to discuss some additional predictions of this approach for the occurrence (or rather non-occurrence) of neutral tone, and for the degree of freedom allowed the poet in deviating from the canonical patterns. Lastly, we give a precise formulation of the rules of tone assignment which turn out to make use of the same labelling conventions already justified for stress trees in section 2.1.

Consider the recitation rhythms shown in (2) above, the (c) versions. If we call the stronger beats S and the weaker W, we may divide the seven-syllabled line into feet as shown:

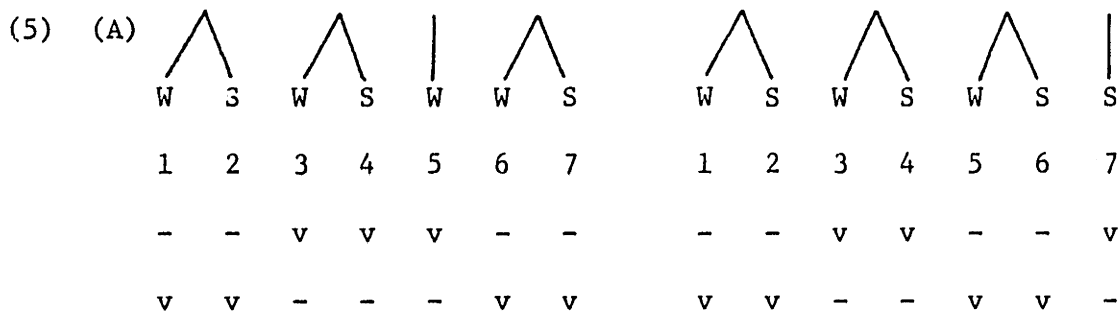


Now if we compare this with the tonal patterns given in (1) and partially repeated in (5) below we notice that, in the case of the first two feet at least, it is the case that two syllables in the same foot share the same tone (a fact undoubtedly related to the restriction to one tone per foot word-internally in the modern language; see section 2.1). Downer and Graham say (1963: 46):

"...the free syllables borrow their tone from the immediately succeeding syllables, the first syllable from the second and

the third from the fourth."

Under an approach like theirs this is an unexplained fact; for us it is a natural result of the division of the line into feet, which constitute the domain of tone. So far we have said nothing about the last three positions, which do not share a single tone. Notice however that the remaining recitation patterns (2a) and (2b) would suggest two different divisions into feet; if we then continue to maintain that all syllables in a single foot share the same tone we will correctly account for the observed tonal patterns. Let us summarize the above:

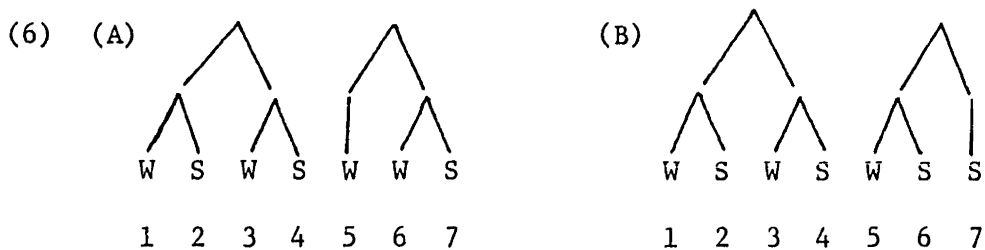


(The choice of W,S for the unpaired syllables is the result of other levels of structure, as will become clear later.) Of course, so far other tonal patterns are still possible; for example, the first two feet apparently always have different tones, and so far this is unexplained. But already the assumption of an underlying foot structure has unified previously coincidental facts. Note that the same arguments hold for the five-syllable line: for economy's sake only one type of line will be used to illustrate each argument except where they differ.

Now let us add in a third consequence. The free positions were

summarized by Wang Li (1957) as follows: "Licence for 1, 3 and 5 strictness for 2, 4 and 6." Notice that these free positions are always foot-initial, and therefore weak. This is not surprising; if strong positions are more prominent, departures from the norm would tend to disturb the pattern more than departures in less prominent (i.e., weak) positions.

Lastly, let us show how this foot structure matches the syntax. We are justified from the recitation facts, which clearly group the fifth syllable with the last two and interpose the major caesura between positions four and five, in further grouping the feet into pairs to give two half-lines:



In the vast majority of lines the syntax conforms to this overall pattern, with the major syntactic division coinciding with the half-line as defined here (see Chen 1979: 406ff for a detailed exposition).

We can now see that the simple division into feet on the basis of recitation data also accounts immediately for some of the restrictions on tone patterns, which positions are weak, and the syntax-metrics mapping. Instead of learning a collection of unconnected facts which could equally well have been different, only a single structure is now necessary and all else follows.

I shall now move on to the second aim of this paper: to advance two further arguments for the metrical structure of regulated verse.

In so doing I shall have cause to spell out that structure in greater detail than has so far been given, and insofar as space permits I will justify each refinement.

2.2.2 Neutral Tone in Min

Boyce (1980) cites some interesting data from Southern Min dialects on verse recitation, and Chen (1980) shows convincingly that a metrical analysis explains the occurrence of combination or isolation tones in different positions. Essentially, isolation tones occur foot-finally and combination tones foot-initially or medially. So the seven-syllabled pattern is usually:

CI / CI / CCI

Where deviations occur they coincide with the predictions of the higher levels of structure, the last syllable is always I, and the end of the first half-line is nearly always I, but the end of the first foot may be C:

CC / CI / CCI

Further, metrical override syntactic considerations in conditioning tone sandhi. I should like to discuss one particular aspect of Min recitation: the relative rarity of neutral tone by comparison with normal speech. In what follows C = combination tone, I = isolation tone, N = neutral tone. Chen gives three cases where there are more neutral toned syllables in speech than in verse (Chen's (10)):

(7) a)	bo	ɕŋ	lai
	not	return	come
Normal:	C	I	N
Verse:	C	C	I
b)	uat	dzip	lai
	turn	in	come
Normal:	I	N	N
Verse:	C	I	N
c)	si	k'i	
	die	away	
Normal:	I	N	
Verse:	C	I	

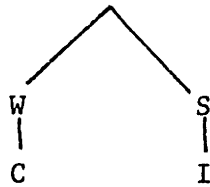
Chen says:

"It seems that in verse recitation the speaker tends to avoid neutral tones wherever possible perhaps in order to maintain the isochronic flow of poetic rhythm, as neutral tones are short and staccato, tending to introduce hiatus in the vocal rendition."

I should like to suggest that we can say something much stronger.

Neutral toned syllables are unstressed and weak; foot-final position is strong and stressed. Since neutral toned syllables would occur in foot-final position in the above examples, it would be necessary to re-label the tree for them to occur. To put it another way, the normal matching for an iambic foot is:

(8)



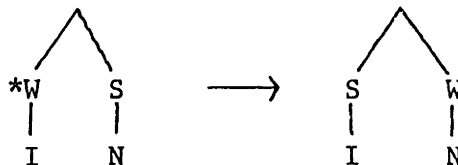
As we saw in section 2.1, neutral toned syllables in Mandarin and Min are metrically unbranching and cannot occur as the DTE of a foot. To put it another way, neutral tone requires a weak node: the configuration:

(9)



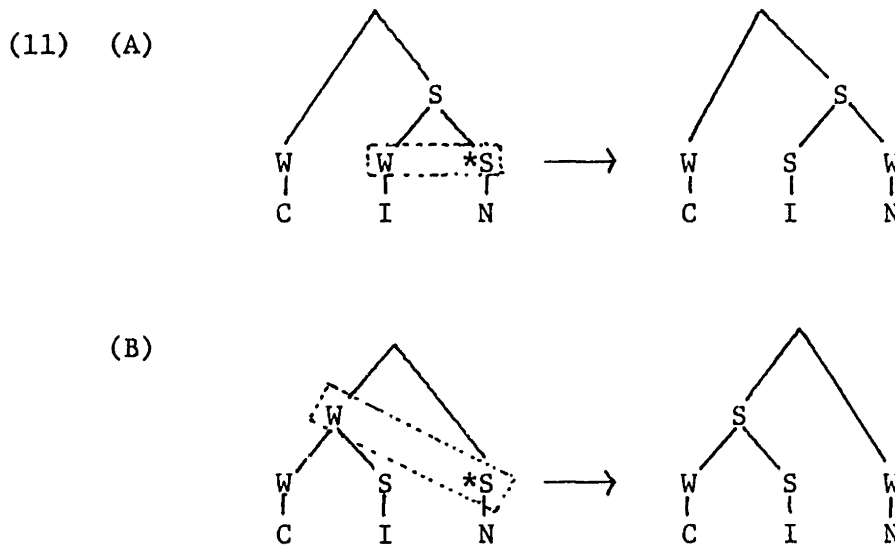
is ill-formed, just as in English a strong may not dominate a syllable marked [-stress] (see L&P:1977). So in order to get a foot-final neutral the ill-formed tree must be re-labelled as shown in (10):

(10)

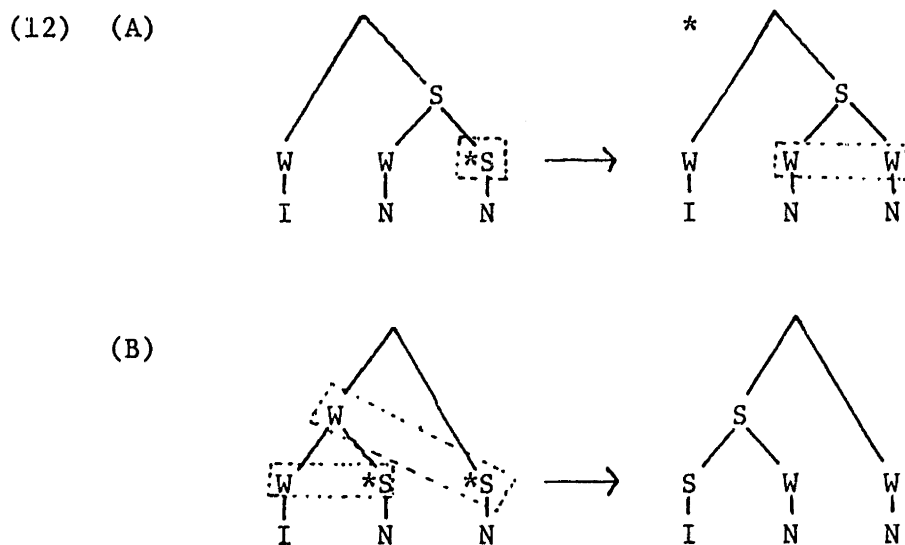


But this is metrically highly marked: a trochee among iambs. It will therefore be avoided whenever possible, but can occur on occasion -- in a different line (Chen's (10a)) si k'i is read I-N, not C.I.

The picture is even more interesting when we consider the three syllables of the second half-line, for which the normal rendition is CCI. Notice the re-labelling required to allow a final neutral tone in the two different structures proposed in (6) above:



((11) assumes that higher levels of the tree are also labelled W-S, like the feet. We will see more consequences of this assumption later.) So we would expect from (11) that a single neutral tone should be possible, but highly marked; one such example is given in (7b). But would we ever expect to find a sequence of two neutral tones? The necessary re-labellings are given in (12):



A metrical analysis would predict that whereas the highly marked (12B)

might conceivably be found, (12A) is impossible since it violates the most basic property of metrical trees: that two sister nodes always hold opposite values. A W-W foot is thus impossible.⁵

We see then that the natural assumption that neutral tone must be dominated by weak accounts for its rarity in verse, and further predicts the extreme rarity of sequences of two neutral tones.

2.2.3 Deviations From The Canon: The Free Positions

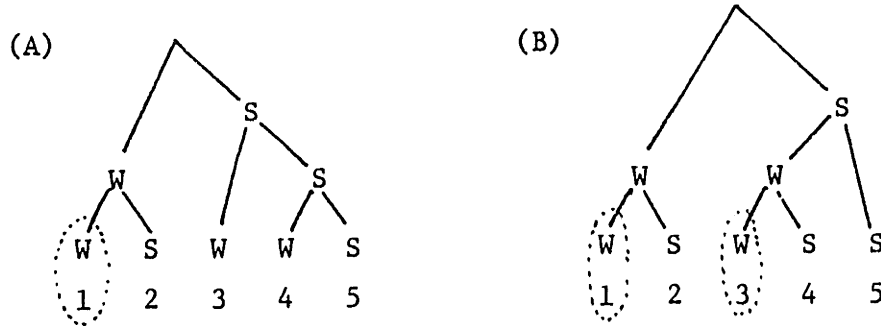
It was pointed out earlier that the free positions, being odd-numbered, are always foot-initial and weak, which would explain their freedom. However, it is not true that all weak positions are free, or even that all those which are free are equally free. How then can we characterize the degrees of freedom? Chen (1979: 397) proposed an algorithm for assigning numerical values to each W and S node, from which the relative weakness could be computed. Anything with an index of -2 or lower is free. I should like to suggest an alternative approach.

The essence of metrical trees is their relative nature. Any two sister nodes have opposite values w-S, and if one changes so does the other. Further, what the labels W-S mean is just that: node N_1 is weaker than node N_2 , and so on throughout the tree. A logical conglomeration of such relative weakness should then tell us which is the weakest, which a little stronger, and so forth: a hierarchy of weakness, in fact. There seems to be no strong reason for assigning numerical values, which obscure the relative nature of the phenomenon.

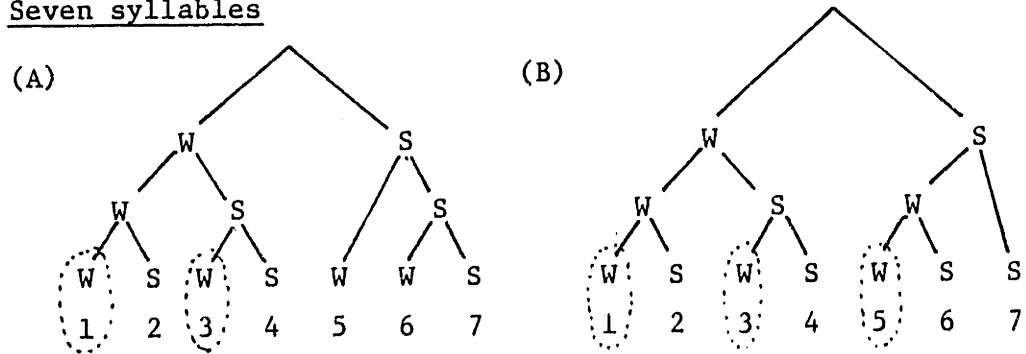
Below the free positions are circled. Notice particularly that

pattern B has one more free position than pattern A, for each line length.

(13) Five syllables



(14) Seven syllables



Notice first that the (A) patterns each contain two W positions that are not free; to be dominated by weak is therefore a necessary but not a sufficient condition for freedom. It is apparently also necessary for there to be a further weak node dominating a position before it is free. The full conditions may then be stated as follows:

Any position that occupies the weak position in a foot
that itself occupies a weak position in the line is free.

Notice that 'weak position in the line' must be interpreted to mean 'weak at any level in the line' (contrast (14A) position 3 and (14B) position 5, both free). Notice also that the phrase 'weak position

in the foot' is undefined for a non-branching foot such as (14A) position 5, which can thus never be free. We also need to explain why position one is special in that it is always free: obviously this is because it is the only position dominated entirely by weak nodes (exactly as in English main stress falls on the syllable dominated entirely by strong nodes). Notice that Chen's numerical algorithm needs two arbitrary cut-off points, one dividing free from non-free positions, and one dividing position one from the other free positions. No such mechanisms are needed under my proposal.

The above provides a starting point for what follows: I shall now proceed to suggest that, contrary to the traditional dictum which considers only the positions circled in (13) and (14) to be free, all positions dominated by a terminal weak may be somewhat free. The argument will concern only the five-syllable line; whether or not it holds true for the seven-syllable line I do not know, although of course I would predict that it would.

Bodman (1978) has cast doubt on the reality of some aspects of the patterns discussed here, since large numbers of lines deviate from the expected patterns. Notice that (see figure (15)) in a five-syllable line positions 2 and 4 are always different, as are positions 3 and 5; yet in many actual cases we find 3=5, and 2=4. Given that in pattern B position 3 is free, we would expect a percentage of lines with positions 3 and 5 the same: - - - v - or v v v - v or the equivalents with position one also changed: v - - v - or - v v - v. Examples of such lines are given below. Numbers are from the edition of Yù Shǒu-Zhēn (1965), poem number followed by line

number.

(16) v v v - v Couplet from 108, 1.7-8:

- - - v - Cháng Jiàn 'At a Buddhist Retreat Behind Po Shan Temple'

'The myriad noises are all hushed now

And I hear only the toll of the bells.'

萬 籟 此 俱 寂

wàn lài cǐ jù jì

惟 聞 鐘 磬 音

wéi wén zhōng qìng yīn

v - - v - Two non-consecutive lines from 117, 1.2, 5

- v v - v Mèng Hǎo-rán 'Feelings at the beginning of winter'

'A north wind blows cold on the river'

'My homesick tears are spent in a land of strangers'

北 風 江 上 寒

běi fēng jiāng shàng hán

鄉 淚 客 中 盡

xiāng lèi kè zhōng jìn

However, there are also unexpected lines which appear to be the result of a free position 3 in a type A line: - - v v v or v v - - -, or with position one also free: v - v v v or - v - - -. Cháng Jiàn's poem (highly deviant, as you can see!), provides one example of each type:

(17) - - v v v 'At first light I entered the ancient temple' 1.1

清 晨 入 古 寺
qīng chén rù gǔ shì

- v - - - 'Shadows in the deep pool clear our minds' 1.6

潭 影 空 人 心
tán yǐng kōng rén xīn

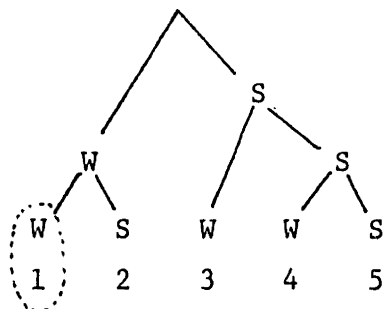
A more substantial problem is apparently posed by Bodman's second category of exceptions: poems with lines in which positions 2 and 4 are the same. Since according to the traditional definition neither position is free, there is no source for such lines. What is more, they are fairly common. Bodman states (p. 109):

"Out of eighty poems, thirty-five contained at least one line in which second and fourth syllables were in the same tone; and a full sixty contained at least one line in which syllables three and five were in the same tone."

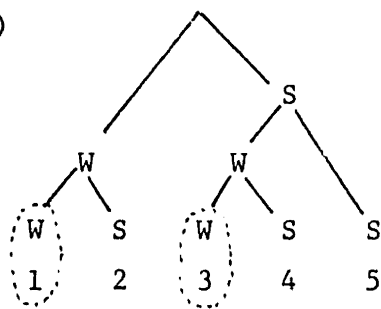
Bodman's sample was the 80 poems in five-syllable regulated verse in Yù Shǒu-Zhēn (1965). Further checking of the same sample revealed some interesting facts. Nearly all the cases of 2=4 were in lines where 3=5; the reverse did not hold. Only four lines of verse (out of 640) had 2=4 but 3≠5. The freedom of the second or fourth position is thus apparently entirely dependent on the freedom of the third position. In terms of percentages of numbers of lines, rather than numbers of poems with deviant lines, about 17% have 3=5 but only 8% have 2=4, and this 8% is a subset of the 17%.

Let us look back at the two tree configurations:

(18) (A)



(B)



Circles mark positions already identified as free. We know from lines such as the last two cited in (17) that A3 can also be free under some circumstances. Suppose then that all terminal weak have some freedom, including the fourth position in tree A. This would give lines like: v v - v - and - - v - v with both positions 3 and 4 being free. These are the lines which make up the deviant 8% where 3=5 and 2=4. For example:

(19) - - v - v Wang Wei 'On missing my way to the monastery of Heaped Fragrance'

'The voice of the torrent gulps over jagged stones' 1.5

泉 聲 咽 危 石

Quán shēng yàn wēi shí

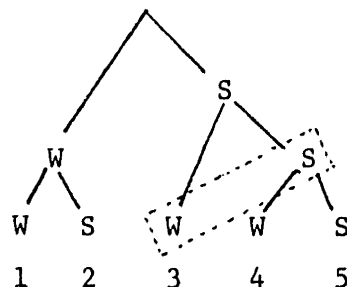
Notice that position four in type B lines is not free. If it were, and the same implication held (4 free only when 3 free) we would find lines of the type - - - - - or v v v v v. Such lines do exist, but are very rare and very highly marked. A table showing the occurrence and non-occurrence of deviant lines appears below.

(15)

	3 free	1 and 3 free	3 and 4 free
A	Not predicted but occurring		
	- - - v v	- - v v v	v - v v v
	- - v - v	- - - - -	- - - - -
	v v v - -	v v - - -	- v - - -
	v v - v -	v v - v -	v v - v -
B	Predicted	Not predicted and non-occurring	
	- - v v -	- - - v -	v - - v -
	- - - - -	- - - - -	- - - - -
	v v - - v	v v v - v	- v v - v
	v v v v v	v v v v v	v v v v v

To summarize, then, the majority of lines, over 80%, do obey the canonical forms in these respects (they may have position 1 free however). Positions which are inherently weak (i.e., the weak half of an iambic foot) and also occupy a weak position in the line tree are free. Other terminal weak are partially free, but subject to a strict hierarchy which only permits the less weak (i.e., position 4) to be free when the weaker (position 3) is also free. The relative weakness is defined because there is a level of sister nodes at which 3 is dominated by W and 4 by S:

(20)



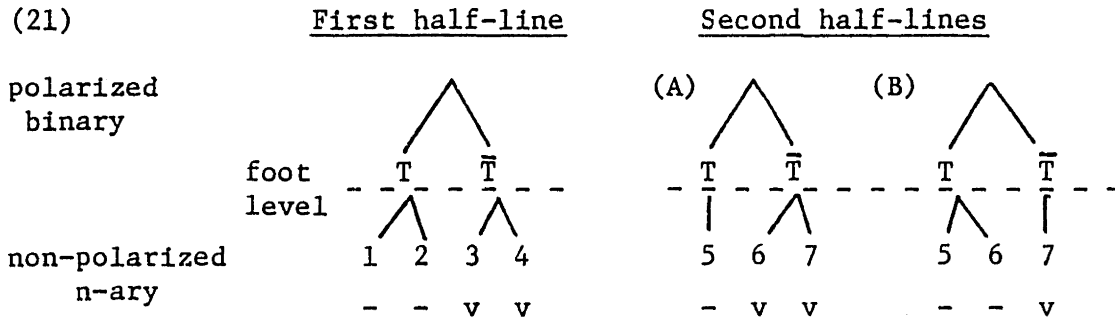
It is worth mentioning, I think, that Bodman takes the small number of complete poems with no violations of the 3≠5, 2≠4 principles to be evidence that such principles do not play a central part in the verse. Since deviant lines are in fact a small percentage of the total number of lines, however (especially if one excludes whole poems made up of deviant lines) and the ways in which the lines deviate is non-random, I still consider the principles to be real.

2.2.4 Tone Assignment

The above sections have, I hope, shown clearly that the foot structure originally motivated by the recitation rhythms also explains a number of other facts about the degrees of freedom permitted the poet. I should now like to return to the question of how this same structure allows a simple formulation of the permitted tone patterns.

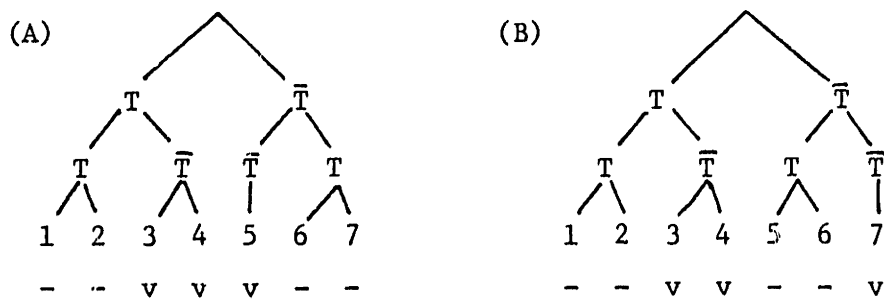
Consider the seven-syllable line shown in (5) above. It is easy to see in the case of the first two feet that whereas syllables within a foot share the same tone, two sister feet must have different tones. The foot level, then, is obviously distinct, and we need only label each foot for its tone and allow this to percolate down to each syllable (see Halle and Vergnaud 1978 for other examples of percolation. Trees of this type may be called non-polarized since sister nodes have the same value of the relevant feature; further, such trees are multi-branching rather than binary). Further, above the foot level the labelling obviously follows the principle of opposition common to all polarized, binary metrical trees: two sister nodes always carry opposite values of a particular feature. We will follow Chen in using T, \bar{T} as variables for the tones even and oblique;

ultimately each may be expanded as either, giving two different realizations of each structure. Let us summarize the above on a diagram of the half-lines:



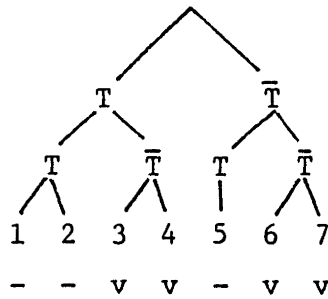
So far, then, we have accounted for the tonal patterns within each half-line by adding the simple assumption that feet, not positions, are labelled for tone. All else follows from the universal properties of metrical trees. However, there are also constraints on which half-lines may go together. Below we give the permitted lines, in (22), followed by those which are not permitted, in (23):

(22) Permitted lines

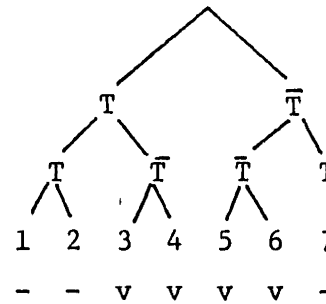


(23) Not permitted

(A)



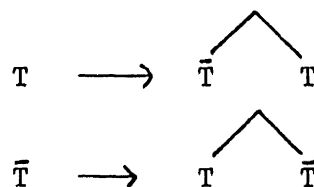
(B)



Up until this point nothing I have said concerning tone assignment is incompatible with Chen's approach (as detailed in Chen 1979), although the line of argument has been somewhat different. From this point on, however, I shall suggest a quite different solution. Let me first summarize Chen's approach, noting where it seems to me to be unnecessarily cumbersome. Then I will propose an alternative.

Chen gives the following rules for tone assignment: (1979: 382)

1. TONE ASSIGNMENT: Opposite tones (T and \bar{T}) are assigned to sister constituents down to the level of the metrical foot in this fashion:



2. TONE SPECIFICATION: T may assume the value of either E (even) or O (oblique), and \bar{T} is opposite to T , subject to the TONOTACTIC CONDITION.

The tonotactic condition is an alpha-switching operation, needed to block the occurrence of lines containing a sequence of four identical tones. Chen gives the condition as follows:

3. TONOTACTIC CONDITION: if TONE ASSIGNMENT produces four consecutive

syllables carrying an identical tone, the tones of the second half-line undergo alpha-switching (E to 0 and vice-versa).

Consider a tree whose highest nodes are labelled $T\bar{T}$ respectively by Tone Assignment (the same arguments will of course hold for trees starting with $\bar{T}T$, which will produce a duplicate output). The expansion rules will then produce trees which are formally identical to those in (22A) and (23B) (except trivially that the feet are labelled $\bar{T}T\bar{T}$ instead of $T\bar{T}\bar{T}$: the specification rules will merge the two cases). The tonotactic condition will then convert the unacceptable (23B) into the acceptable (22B), which is not independently generated. This seems extremely suspicious: the basic rules generate half good and half bad lines, and the tonotactic condition then fixes things up. The problem arises because Chen's tone assignment rules assign the same tones to the two different trees underlying patterns A and B. If they can be assigned different tones directly, the need for the condition will disappear.

Let us explore the consequences of referring to the different geometry of the trees underlying A and B in the tone assignment rule. The generalization seems to be that, in the second half-line, the branching node is labelled T. We might state this as follows:

- (24) In a pair of sister nodes N_1, N_2 , N_1 is labelled T if and only if it branches.

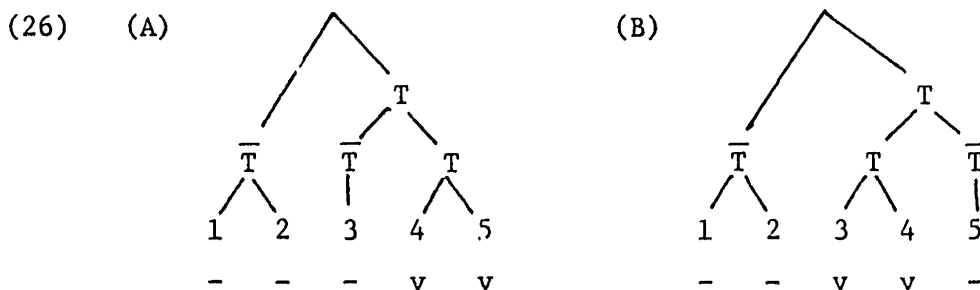
It can easily be seen that this will generate all and only the permitted trees in (22). The fact that the restriction may be stated

in this form is of particular interest because it has been proposed (Halle and Vergnaud 1978) that metrical trees in all languages are labelled W-S by one of two conventions, (24), or its mirror image (25):

- (25) In a pair of sister nodes N_1, N_2 , N_2 is labelled T
(or S, in a stress tree) if and only if it branches.

The child has only to discover which of these conventions is at use at that level in his language. English stress, for example, uses (25) for W-S labelling, and so does Mandarin foot stress, while Mandarin word stress uses (24) (see 2.1.1).

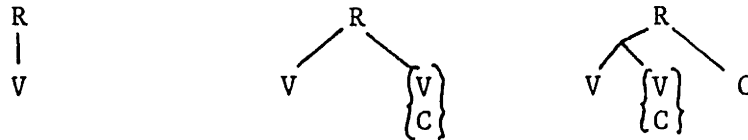
This brings us conveniently to the last problem: tone assignment in the five-syllable line. Suppose that the five-syllable line makes use of convention (25). We will then get the following lines:



These are of course exactly the permissible five-syllable lines. Chen had rejected a proposal using (24) on the grounds that a tonotactic condition would still be necessary for the five-syllable line, but we have seen that this is not the case: instead the other universal labelling convention operates in the five-syllable line.⁶

Paul Kiparsky has suggested a different approach that allows a unified labelling principle for both line lengths. The idea is that distinctions can be made not only between branching versus non-branching,

but also between the degree of branching exhibited by a node. So some languages distinguish between light, heavy and superheavy syllables in stress assignment, and this distinction can be viewed as the distinction between non-branching, branching, and 'super-branching' rhymes:

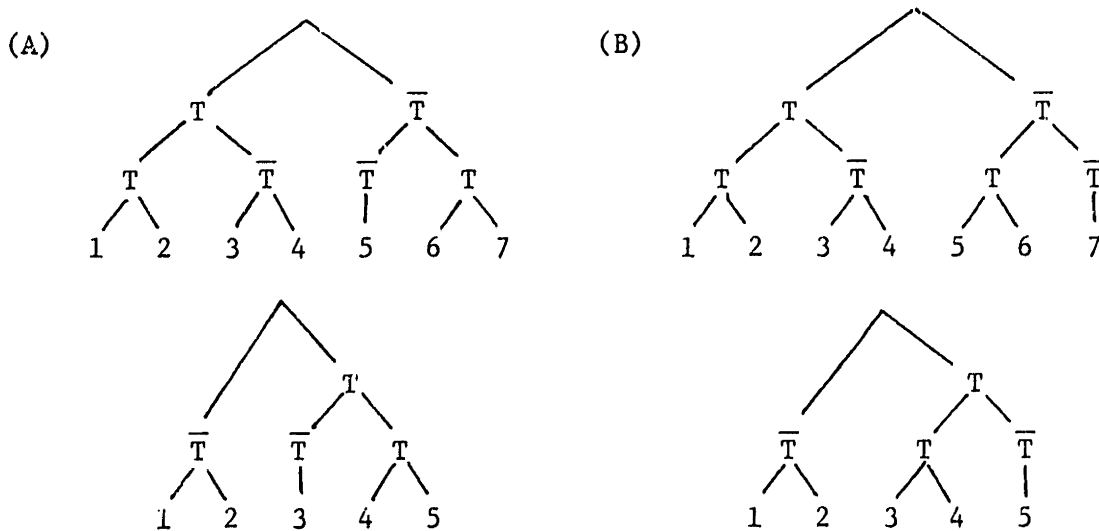


If geometric discrimination of this kind is admitted, the tone patterns of lù shī can be accounted for by a single rule:

(27) Label N_2 T iff stronger than N_1

where 'stronger' is defined as 'more branching' (or 'dominating more terminal nodes'). When both nodes are identical the first will be labelled T.

This will apply to the four possible trees as follows:



thus generating all and only the desired forms.

Notice by the way that there is no real evidence for any labels above the foot level; it may well be that only feet are labelled for tone, in which case the same labelling conventions would still be in operation, but they would stop operating once all feet were labelled. Since Chen labels from the top down, under his approach the entire tree must be labelled.

Some rather abstract notions have been introduced here to account for the facts of regulated verse. Most of this machinery has been found to operate in other languages in metrical phenomena, and has been proposed to be universal. Its apparent use in tonal patterning is therefore of great interest. Notice how natural this form of verse becomes under the assumptions made here. Regulated verse has lines made up of iambic feet, of which the weak positions are free. The feet are labelled for tone by one of the two universal labelling conventions (24) and (25), and the poet must only learn which convention applies to which verse form. All else follows.

NOTES TO CHAPTER TWO

¹ In section 5.1 we will argue that in fact Mandarin has only phrase stress, and word stress does not exist as a separate level. This does not affect the discussion here.

² Monique Roi (personal communication) had pointed out to me that there is also a large class of words usually ignored in the literature with the stress pattern 1 2: Zhōngguó 中国 'China'. Under any approach these must be lexically marked somehow, presumably with a W on the second foot given the theory expounded here.

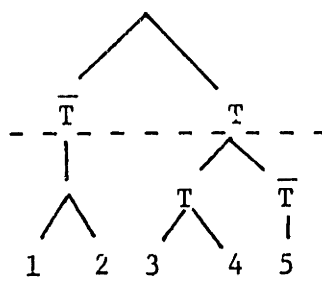
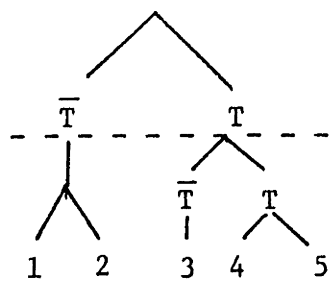
³ Apparently it is also possible for some speakers to have a full-toned e¹³ with either pattern; this is unexpected, and I have nothing insightful to say. Paul Kiparsky has suggested to me that this might be the result of restressing at a late stage in the derivation.

⁴ This section is a revised version of Yip (1980). In the same volume are a number of interesting papers on the same topic.

⁵ Matthew Chen (personal communication) has pointed out that this also predicts that in seven-syllabled lines we should never find two neutral tones in positions 3 and 4. For what it is worth this prediction appears to be true, and in fact Chen found no such sequences anywhere in the data he checked.

⁶ James Huang has reminded me that if one takes the half-line level seriously both seven and five syllabled lines can be dealt with by (24), since in the five syllabled line the first half-line never

branches, and it will always be labelled \bar{T} as desired:



3. AN INTRODUCTION TO THE CONCEPT OF REGISTER

3.1 A Summary

I have referred from time to time to the 'register' of a tone without much explanation. In this chapter I will explain exactly what is meant by this term in the present context, and give arguments for its existence from two main perspectives -- the typological and the historical. In the case of Chinese a detailed argument will be developed tracing the tonal evolution from Middle Chinese to early Mandarin and Cantonese. We will also have a few words to say about the physiological correlates of the feature Register.

It is first necessary to clarify the terminology and formalism that will be used from now on. It is maintained that the pitch of a syllable is the phonetic manifestation of a phonological tone that is itself made up of a two-feature complex. These two features are Register, which may be [+ Upper], and the melody or Tone, which may be [+ High] (usually abbreviated to H and L). The reader should be aware that the word 'tone' is being used to refer to two different things; with an upper case T, Tone, it refers to a single binary feature, but with a lower case t, tone, it refers to the entire phonological representation that determines the phonetic pitch, including both Tone and Register. When the possibility of confusion is high I will try to remind the reader of the sense intended; at times the term melody may be used to refer to a Tone or sequence of Tones. Where none of these distinctions is relevant the word tone may be used loosely to refer to surface pitch, as in: yip²² has the tone 22.

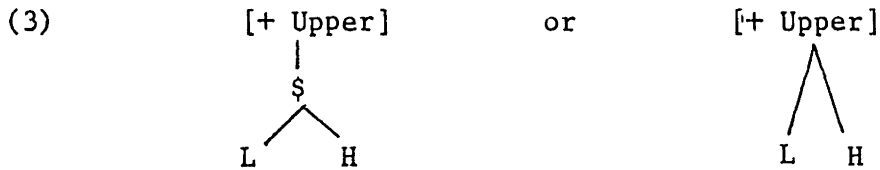
These two features Register and Tone interact to define four pitch levels in the following way. Register is 'dominant' and splits the pitch range of the voice into two halves. Tone then further subdivides each half to give a total of four levels:

(1)	<u>Register</u>		<u>Tone</u>
			+ High (H)
	+ Upper		- High (L)
			+ High (H)
	- Upper		- High (L)

Chinese languages normally make rather full use of this system, but some languages may only use one or other contrast. It is further maintained that each of these features is potentially an independent autosegmental tier and therefore subject to the WFC. However, only the feature Tone may occur in sequences underlyingly: Register remains constant over the morpheme. The main effect of this is to restrict the basic inventory of tones to no more than two of any given contour:

(2)			L H
	+ Upper		H L
			L H
	- Upper		H L

We will see in section 3.2 that this is a highly desirable consequence. Since both Register and Tone form autosegmental tiers we will encounter representations like the following:



The second representation does not include the segmental level but simply associates the two autosegmental tiers with each other. Since registers do not occur in sequence the two are equivalent, and the second will often be used for brevity.

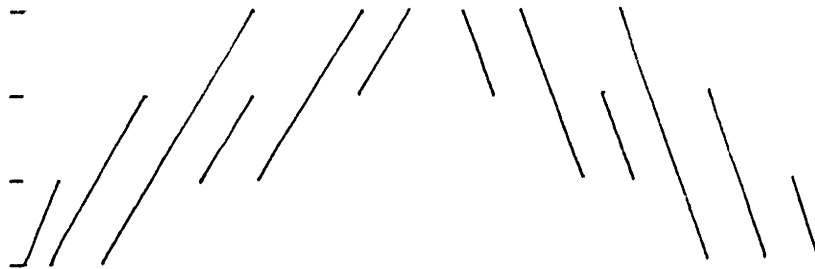
Note by the way that the term Register is used in a number of different senses elsewhere in the literature. In S.E. Asia it commonly refers to the quality of phonation, such as breathiness or creakiness, rather than to absolute pitch. No such sense is intended in most languages here, although we will see that just such a historical source may be postulated for the purely pitch manifestations of Register that concern us here. Goldsmith uses it to refer to the level on which a High or Low tone is realized, so that the High or Low 'pitch register' can be moved up or, more commonly, down (see Goldsmith 1976a: 104-110). It may be possible to maintain that this concept of register in fact makes use of the same feature Register that is argued for here on different grounds, and we will discuss this in section 3.2.3.

In the next section we will give some typological evidence that supports the feature system laid out here over earlier proposals.

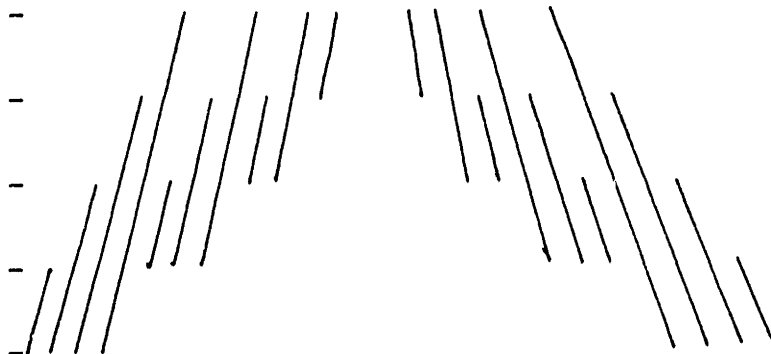
3.2 Typological Considerations

Previous feature systems for tone have usually assumed that it is necessary to represent four or five pitch levels and a variety of contour tones. Assuming that contour tones are made up of sequences of level tones all such systems (Woo 1969, Maddieson 1970) share a common defect: once four levels are defined it is possible to represent six rising and six falling tones by combining the levels, and with five levels there are ten rising and ten falling tones possible. This is a gross over-generation since no language has anything like that number of underlying contrasts and the vast majority have no more than two rising or two falling tones. Below we show the possible contours given four and five levels:

(4) Four levels:



Five levels:



The same problem arises in Wang's system (Wang 1967). In his system five pitch levels are defined by three features: High, Central and Mid. He then notes (p. 99):

"We are only allowing two pitch levels to be distinctive for [+ Contour] tones...Such restrictions can always be removed when counter-evidence becomes available."

He achieves this by means of a redundancy rule:

$$(5) \quad [+ \text{Contour}] \longrightarrow \left[\begin{array}{l} - \text{Central} \\ - \text{Mid} \end{array} \right]$$

thus allowing only two rising and two falling tones which are [+ High] respectively. The fact remains that without this restriction Wang's system predicts five rising and five falling tones, and the restriction is in no sense a natural part of the system but rather an extra constraint imposed upon it. As Wang points out, this would prove an advantage should evidence of a language with more than two contour tones of the same shape be found, but we will discuss the few such cases that have been reported in the literature and show that in no case are they very convincing. If it is the case that no such language exists, it would be desirable for this to be an automatic consequence of the feature system, rather than something requiring an extra-systematic statement.

The feature system being proposed here has exactly that effect; since Register must remain constant over the morpheme but Tone can change there is the possibility of only one contour tone in each register, as shown in (2). Notice that we are not claiming that the maximum number of contrasts of a single contour at the phonetic level

is two; it is quite possible that a greater number of contrasts might arise as the result of morphological and phonological processes of various kinds. For example, Mazateco as described by Pike (1948: 95ff) has four level tones and a wide variety of contour tones on the surface. Using the Mexicanists system in which 4 denotes the lowest pitch and 1 the highest pitch there are on the surface at least the following contour tones on single vowels:

<u>Rising</u>	ʃki ⁴³	'medicine'
	ti ⁴²	'bowl'
	the ³²	'our (inclusive) foreheads'
<u>Falling</u>	ti ³ hnti ¹³	'dirty boy'
	ni ³ tha ³ hma ²³	'black griddle'
	ntʔia ³ va ³ se ³⁴	'town hall'
	vha ³ htɿ ¹⁴	'we (exclusive) tighten (Timeless)'
	vʔe ² ʃi ²⁴	'we (exclusive) dry (Timeless)'

However, Pike says (p. 97):

"We have already shown that noun stems may occur with any of the level tonemes. In addition, however, there are two lexical toneme combinations that appear on such stems. One of these is a glide from pitch level 4 to pitch level 3, and the other is a glide from 4 to 2. Other combinations have not yet been found."

The same is true of verb stems (see Pike, p. 107). That is, two rising tones can appear underlyingly, but all other contour tones arise only as the result of morphophonemic processes. For example, the other rising tone 32 is formed as follows:

thɛ³ + a² → thɛ³² 'forehead + 1st pl. incl. possessive pronoun'

And some of the falling tones are formed as follows:

va¹ + htɪ¹ + i⁴ — vha³htɪ¹⁴ 'he places + a pile + 1st pl. excl.'

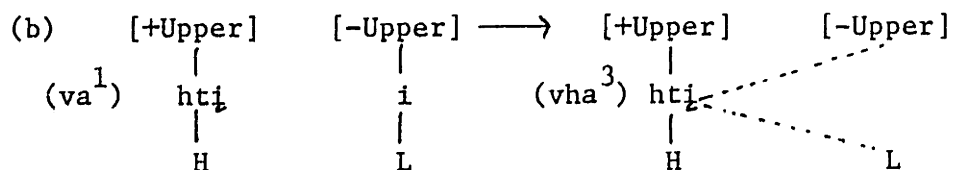
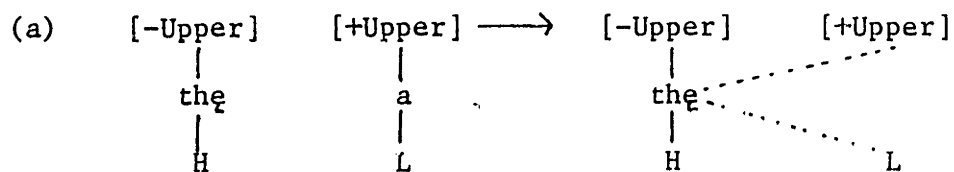
vʔe¹ + ši² + i⁴ — vʔe²ši²⁴ 'he deposits + dry + 1st pl. excl.'

(The changes in the segments and tone of the first syllable are irrelevant here.)

If we now try to capture this in the feature system proposed here, the underlying tones are:

- 1 [+ Upper, H]
- 2 [+ Upper, L]
- 3 [- Upper, H]
- 4 [- Upper, L]
- 42 [+ Upper, LH]
- 43 [- Upper, LH]

and the derived tones are the result of multiple association after vowel contraction:



There is thus no problem in simultaneously constraining the system so as to restrict the underlying inventory and describing a wide number of surface contrasts within the same contour shape. I take this to be a major advantage of the system.

The rest of this section will be concerned with discussing languages that cannot be described by the system, and whether or not they require its revision. We will conclude that the data is in all cases too sketchy to justify weakening the system to permit a greater number of underlying contrasts than the four level tones and two of each contour currently predicted. Rather we take it to be more important that this system correctly characterizes the overwhelming majority of tone languages that have been carefully described. C-C Cheng has collected data on about 900 Chinese dialects, and in no case does the tonal inventory exceed the possibilities predicted by this theory.

Since the feature system proposed here can describe four level tones and two tones of any given shape, counter-examples would be of two kinds:

- (i) languages with more than four level tones
- (ii) languages with three or more tones of a single shape

Let us take these in turn.

3.2.1 Five-Tone Languages

Languages with two, three or four level tones are quite common, but there are very few reports of languages with five or more level tones. By far the most convincing case is Miao-Yao (Chang 1953).

Tahua Yao and Hei Miao are both reported to have five level tones (in addition to different inventories of contour tones), and although these are to some extent predictable from the initial consonant they are apparently in true contrast in some environments. Unfortunately there are no data available on any tone sandhi or morphological tone changes that might help make it clear whether these tones are all phonologically level as well as phonetically level, but it would appear to be a true case of five level tones, and we will return at the end of this section to how to deal with it.

The second candidate for five level tones is an African language, Dan (Bearth and Zemp 1967). However, this case is not as convincing as the Miao-Yao facts, because careful reading of their report shows that one of the five level tones is not in fact level in the actual frame they used to set their pitch levels, or in the majority of environments. It would seem then that the language in fact has four level and one falling tone, and does not constitute a counter-example.

The remaining cases are very much less convincing; the existence of five phonetic levels does not in itself show the need for five phonological levels, and only if the phonological rules themselves need to make reference to five levels of tone is it necessary to provide for a five-way distinction in the feature system. For the other cases of five phonetic levels it appears to be possible to deal with the phonology in terms of a four-way distinction only, as pointed out by Wang (1967) for Trique (Longacre 1952), and Anderson (1978: 146) for Ngamambo Bamileke (Asongwed and Hyman 1976). For references to the remaining cases see Maddieson (1979). Note that I have not yet

confronted the problem of specifying five phonetic levels, as is certainly necessary. I take the position that the phonetic instructions may require a further elaboration of the phonological features, perhaps from binary to n-ary, but will not discuss the details here. Admitting all such phonetic distinctions into the phonological features grossly overgenerates tonal inventories, and is therefore a serious error.

There are two conclusions one can draw from the above data:

- (i) for the vast majority of languages four is the maximum number of contrastive level tones.
- (ii) a fifth level may be possible in highly marked cases, although this is still in doubt.

This suggests that in the vast majority of cases the feature system should not provide for more than a four-way contrast, and that if a fifth is needed it should be achieved by addition of an extra feature to the system. Previous approaches have on the contrary assumed that the rare need for a five-way contrast should be allowed to determine the features in general use, but as we have seen this results in an undesirable overgeneration of contour tones as a side effect. As to how the feature system presented here might be extended to deal with a fifth level, one can only speculate in the absence of phonological information about such languages. Presumably a feature akin to Woo's [± Modify] might be introduced, or it might be that the fifth level is the surface manifestation of a toneless syllable.

3.2.2 Three Contour Tone Languages

The second set of counter-examples involves languages with more

than three tones of a single contour. If such languages exist they are very rare, like the five-level tone languages, but a few cases have been reported. The best known is probably Tepetotutla Chinantec (Westley 1971), which is described as having three rising tones. One of these, we are told, is perceived as 'almost level in tone' and may in fact be phonologically level -- there is not enough information to decide. In any case, all three putative rising tones are only in contrast on a small number of syllables bearing a particular type of stress (termed 'controlled' as opposed to 'ballistic') and not ending in glottal closure. Other syllables may have only two of the three glide types. All in all, then, the language is probably not a counter-example, and the third glide is a stress-conditioned variant of some other tone (i.e., a problem for the phonetics rather than the phonology).

Some S.E. Asian languages have been reported as having three rising tones, including Tonkinese Vietnamese (Barker 1966). Many of these languages have distinctions in voice quality as well as pitch, as in this case one of the three is described as 'high rising broken', and presumably differs in quality (by having medial glottalization or creakiness) rather than just in pitch from the two simple rising glides. Provision for distinctions of this kind must be made in any system, and this is therefore not a counter-example. Interestingly, it is this broken tone that is not distinguished from one of the other glides in most other Vietnamese dialects, which may have two rising tones, suggesting that a three glide system of any type may in fact be highly marked.

Within the Chinese languages the only counter-example of which I

am aware is the Min dialect of Kienyang (Norman 1973), which has the following tonal inventory:

Level:	33			
Rising:	35			
Falling:	53 31	43 32	21	

There is apparently nothing in the phonology that allows one to differentiate the falling from the level tones (i.e., no sandhi), and it seems highly likely that the last three falling tones, which have a rather slight fall, may be phonologically level but surface phonetically with a slight downdrift giving the appearance of falling tones. Kienyang would then have a very symmetrical tonal inventory of four level, two falling and one rising tones. There seems no reason, then, to consider this a counter-example.

Even though none of these cases seems to require allowing for more than two rising and two falling tones (and thereby provides strong evidence in favour of a system which limits the possible inventory to two of each) it might be interesting to speculate on how the system could be extended if a three-way contrast were found to be necessary. The simplest change would be to remove the restriction that Register remains constant over the morpheme underlyingly. This would then make the system equivalent to other four-level systems in allowing up to six tones of a single contour, but would also fail to explain why the full potential of a six-way contrast is apparently never used.

3.2.3 Downstep and Other Matters

Chinese languages normally make rather full use of the available tonal contrasts for lexical purposes, but many languages do not. Most African languages, for example, have less than four level tones, and two tone systems are very common. If there are only two contrastive level tones only one feature is needed to distinguish them, and we might reasonably ask which feature is used in such cases, and whether the unused feature is used in any other way. As to which feature is used, this can normally be settled quite simply by noting whether or not the language allows sequences of tones on a single morpheme; if it does, the tones must be distinguished by the feature \pm High, since only H and L tones occur in underlying sequences. If it does not, we have a system using \pm Upper only.¹ Most African languages fall into the first category: for example, Mende has H and L tones, and permits the following sequences on morphemes:

H pélé́	HL kényà	LHL nyàhá̂
L bèlè̀	LH nìká́	

Such languages also commonly (although not necessarily) exhibit the phenomenon known as downdrift, by which a H tone is realized at a lower level following a L tone. Subsequent H tones will then be at or below this level. It is tempting to suggest that downdrift is a process that affects the so far unexploited feature of Register, but this would be to claim that downdrift will only be found in languages with no more than two tones, and this is false. Languages with three tones and downdrift are attested and include Ga'anda (Newman 1971) and

Yala (Armstrong 1968).²

It seems then that downdrift must be dealt with independently by a concept of resetting that is triggered by a lower tone. What is reset may apparently be either feature. For example, in a two tone H-L system with downdrift the feature [+ High] is subject to resetting after [- High]. But in Ga'anda, which has three tones of which only one, High, is downdrifted after either of the others,³ it is apparently the case that [+ Upper] is reset after [- Upper], since the three tones are most conveniently represented as follows:

high:	[+ Upper, H]
mid:	[- Upper, H]
low:	[- Upper, L]

The resetting rule will then apply to [+ Upper]/[- Upper] _____ and achieve the desired result. The alternative, with mid as [+ Upper, L], would imply a rule resetting H after L, just like in a two-tone system. Languages like Yala which downdrift both High and Mid after any lower tone may reset both [+ Upper] and [+ High]. For interesting recent treatments of the hierarchical resetting involved in downdrift see Clements (1979) and Huang (1979).

Apart from languages which use only one of the two features contrastively there is another common way in which languages 'collapse' the system. Many Chinese languages have only three level tones, and in such cases there is ambiguity as to how to represent the middle tone:

High:		[+ Upper, H]
Mid:		[+ Upper, L]
	or	[- Upper, H]
Low:		[- Upper, L]

We will see that this ambiguity is necessary to explain the different ways in which mid-tones behave in different languages, and even, in some cases, differences within the same language (see section 5.3 on Amoy). This ambiguity, we will argue, also explains why neutralization under sandhi to a mid-tone is much more common than neutralization to any other tone: the neutralization is phonetic but not phonological since both representations of a phonetic mid-tone are involved: [+ Upper, L] and [- Upper, H], which merge phonetically as [33].

One other interesting effect of this ambiguity is in explaining flip-flop. Flip-flop is a process by which in a three-level tone system the high and low tones interchange but the mid-tone remains unchanged. Obviously this must be an alpha-rule, and would be written:

$$\begin{bmatrix} \alpha \text{ Upper} \\ \beta \text{ High} \end{bmatrix} \longrightarrow \begin{bmatrix} - \alpha \text{ Upper} \\ - \beta \text{ High} \end{bmatrix}$$

Notice the effect this will have on the mid-tone:

$$\begin{bmatrix} + \text{ Upper} \\ - \text{ High} \end{bmatrix} \longrightarrow \begin{bmatrix} - \text{ Upper} \\ + \text{ High} \end{bmatrix}$$

In other words it will be left phonetically unchanged. An example of such a phenomenon in Chaozhou is given in Wang (1967: 102) (although it is possible that in this case more careful phonological analysis

will show that this is not a flip-flop rule at all; this is because the flip-flop analysis assumes that the input is the citation form of the tone, but in fact the rule applies before another tone, and should perhaps take the contextual form as input. The contextual forms are low level and low rising, and no flip-flop is any longer involved).

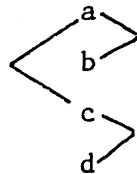
The sort of rules that cannot be stated naturally in this system would be general tone-shift rules that raised all tones but the highest, which lowered:

55	←	[+ Upper, H]	→	[- Upper, L]
44	←	[+ Upper, L]	→	[+ Upper, H]
33	←	[- Upper, H]	→	[+ Upper, L]
11	←	[- Upper, L]	→	[- Upper, H]

This would need two α -rules:

- (1) $\alpha H \longrightarrow -\alpha H$
- (2) $\beta \text{Upper} \longrightarrow -\beta \text{Upper} / \begin{array}{c} \text{---} \\ | \\ \text{L} \end{array}$

While stateable this is hardly elegant, and the absence of such rules can be taken as non-accidental. On the other hand, rules relating non-adjacent tones of certain types can be very simply stated in this system, but are quite hard to deal with in any other. I have in mind alternations like those in Igede (Stahke 1975):



The $a \sim c$ alternation is just $[+ \text{Upper}] \sim [- \text{Upper}]$ in my system (whereas $a \sim b$ and $c \sim d$ are $[+ \text{High}] \sim [- \text{High}]$).

In the next section we will discuss some historical evidence that suggests that Register distinctions arose through a contrast in voice quality (i.e., breathiness) that eventually degenerated into a pitch distinction only. This interacted with the [+ High] distinction already present in some cases to provide a much larger number of tonal contrasts than previously present. The Register system allows a particularly natural characterization of tonal development in these cases.

3.3 Historical Evidence

3.3.1 Why Other Systems Fail

The existing feature systems all run into problems when it comes to explaining the historical development of tone. One of the most widespread changes in tone languages, at least in Asia, was a process of register splitting that effectively doubled the number of tones: every tone present before the split gave rise to two variants, one lower and one higher, but with contours essentially unchanged. The complicated subsequent development of the tones has frequently obscured the picture, but there is little doubt that this is what happened, and in some languages, such as Cantonese, it can be seen quite clearly. The split was conditioned by the properties of the initial consonant of the syllable, and it is generally believed that voiceless consonants gave rise to the higher variant while voiced consonants gave rise to the lower variant (this is oversimplifying, and we will discuss it in more detail later). So in Cantonese we have the following split:

Voiceless initial	53/55	35	44	5
				4
Voiced initial	21	24	33	3

(The extra splits will be discussed below.)

Now consider how such a development is to be dealt with phonologically. It appears likely that an additional binary feature has come into play, and indeed as we shall see this is exactly what can be stated very simply given a register system: the feature [+ Upper] is now being used for tones in addition to the original [+ High].

In the other systems there is no simple way to state the change in terms of the addition of a feature to the system because there is no available feature that has the desired effect. The feature must have the effect of raising/lowering all tones, and it must also be true that without that feature the remaining features can express a simple two-way height contrast. Let us examine the best known feature systems in turn.

Wang's system assigns values for the level tones as follows:

	┌	┐	└	┘	┘
High	+	+	-	-	-
Central	-	+	+	+	-
Mid	-	-	+	-	-

Wang notes (1967: 97) that the feature [Mid] is used only in languages with five levels. We are only concerned here with four levels, so that leaves only High and Central. It seems clear that the basic

system must make use of [+ High], since this is the only feature that divides the whole pitch range into two ([Central] effectively divides it into three). But then the register split would have presumably amounted to the addition of the feature [+ Central], and this unfortunately has the effect of lowering high tones and raising low ones, which is not what happened. Wang's system then fails to account easily for the register split.

Next consider Woo's (1969) system, which characterizes five levels as follows:

	┌	┌	┌	┌	┌
High	+	+	-	-	-
Low	-	-	-	+	+
Modify	-	+	-	+	-

The basic system must have included [High] and/or [Low], and certainly neither of these can have been the feature added at the time of the register split, since between them they only define three levels. [Modify] must then have been the new feature, but this suffers from the same defect as Wang's Central in that it lowers high tones and raises low tones instead of lowering all of them. Woo's system therefore also fails.

Lastly consider Maddieson's approach (1970):

	┌	┌	┌	┌	┌
High	+	+	-	-	-
Low	-	-	-	+	+
Extreme	+	-	-	-	+

The same arguments as for Woo's system apply, with the difference that the added feature, which would have to be [+ Extreme], has the reverse effect of raising the high tones and lowering the low tones. Maddieson's system constitutes no improvement, then, as far as explaining the register development is concerned.

In the rest of this chapter we will try to discover what the source of the register distinction might have been, and will conclude that voicing alone is not sufficient to trigger such a tonal split but that voiced aspiration was necessary. We will propose that an intermediate stage in such a split involves a voice quality distinction in the vowel like the breathy voice in some modern Wu dialects, and that eventually even this distinction was lost, leaving a purely tonal residue. Using Halle and Stevens (1971) features we will suggest that the feature for breathy voice, [+ spread glottis], is automatically linked by convention to a feature [- Upper] that causes low pitch, and that when breathy voice is lost the feature [- Upper] stays behind.

First some background.

It has long been realised that there is interaction between tone and certain consonantal features -- more specifically, that there is a relationship between features such as aspiration and voicing, and tone. Halle and Stevens (1971) in an influential paper proposed a set of laryngeal features which would correctly predict some of these interactions; in particular, they suggested (over-simplifying somewhat) that voicing in obstruents and low pitch in vowels were both caused by a feature [+ slack vocal cords] thus explaining why

voiced obstruents have the effect of lowering the pitch of the following vowel (see Lehiste 1970). This effect has been clearly demonstrated synchronically, although the mechanism is still a matter of dispute. Historically there are many cases where proto-voiced consonants have become de-voiced, but have left the following vowel lower in pitch than those following proto-voiceless obstruents (see Haudricourt 1954 and Matisoff 1973). A re-examination of such cases, however, suggests that something more than voicing is involved. Although it is apparently true that voicing is a necessary condition for the development of two tonal levels where only one existed before, it does not appear to be sufficient. If it were, we would expect to find the following schema:

*p, *ph, *ʔp	—→	Upper Register
*b, *bh, *ʔb	—→	Lower Register

In fact, such cases are rare, if they exist at all.⁴ Rather, register apparently normally develops from voiced aspirates, and even where voiced unaspirates produce low pitch other voiced obstruents (i.e., glottalized or implosive voiced stops) fail to do so. The first part of this section will be concerned with showing that voicing is a necessary but not a sufficient condition for the development of tonal register, drawing evidence from several language families. It will also include cases where voicing is a necessary but not a sufficient condition for changing tonal contour by depressing the start of the tone, in one case in a language not previously tonal (Punjabi). The second part is a detailed examination of the Mandarin and Cantonese cases.

3.3.2 Voicing -- Necessary But Not Sufficient Condition for Register Development⁵

To support my assertion that voicing is a necessary but not a sufficient condition for the development of tonal register I shall introduce evidence from a variety of language families. First, I will discuss several modern Chinese dialects that appear to provide examples of different stages in the process of register development, from Shanghai where register is fully predictable from the presence or absence of 'murmur', to Cangwu, where it has become fully distinctive. Secondly, I will discuss two Indo-Aryan languages, one of which is tonal. Third, I will discuss the Tai case very briefly. Lastly, some short remarks on two African languages, and Tibetan. In the next section I will return to Chinese to discuss the development of the tonal categories in Mandarin and Cantonese in some detail.

(a) Modern Chinese Dialects

Some Chinese dialects, of which the best known are the Wu dialects including Shanghai, still have voiced stops corresponding to those reconstructed for Middle Chinese. In such cases the stops are accompanied by voiced aspiration and breathy vowels (termed 'murmur' by Sherard (1972) following Ladefoged), and low-tone allotones. The nasals exhibit a contrast: if breathy voiced, then low-toned, and vice-versa. So we have ㄊ toq 鴉 but ㄊ dhoq 讀, and ㄋ nɔ 國 but ㄋ nhi 儀. In Shanghai, then, tonal register may be predicted from the presence or absence of voiced aspiration. The connection between the two is so close that when a syllable which begins with a voiced aspirate loses its distinctive tone (e.g., in


non-initial position in a compound) the aspiration is also lost (but not the voicing). Since nasals and liquids may or may not be aspirated, this contrast is neutralized in such positions.

In Kwangsi province there is a group of dialects studied recently by Tsu-ji (1977). The dialects are members of the same general group as Cantonese -- Yue -- but they retain voiced stops and breathy vowels. They illustrate a clear progression from a Shanghai type situation, with register predictable, to a Cantonese type situation, with two phonemic registers in all tones. Rongxian has not only bh and dh, but also b and d. The nasals and liquids have two forms: lh and l, and ngh and ng, for example, and so do the fricatives. All and only those syllables whose initial consonant has voiced aspiration are low toned with breathy vowels.⁶ Cenxi has started to lose aspiration after nasals, so we now find some nasals in the lower register, but without breathy voice. The tone is thus no longer fully predictable. Yulin never shows breathiness after low toned nasals, but in both Yulin and Cenxi the register is still predictable for the obstruents. Nanning Ping-Hua has started to lose aspiration and breathiness in the obstruent-initial syllables too, although where breathiness is still present the syllable is always lower register. And Cangwu has lost all traces in the obstruents too, and shows the same pattern as Cantonese. It is clear that in these dialects the crucial factor in register splitting is voiced aspiration, and not simply voicing.

(b) Voiced Aspiration, Breathy Voice, and Pitch in Indo-Aryan

Indo-Aryan is reconstructed as having had a four way stop

contrast p, ph, b, bh. Gujarati has developed murmured (breathy) vowels from *bh under certain circumstances, and in fact the two are in allophonic distribution in the modern language word finally: labh̄ ~ lab (where v means a breathy vowel). (Note that according to Pandit (1957) Gujarati has no syllables of the form [bh̄ar], but only [bhar], [bar], and [bar] < *bh̄ar .) Fischer-Jørgensen (1967) and Dave (1967) undertook a detailed acoustic analysis of Gujarati vowels, and found that the murmured vowels were significantly lower in pitch than the clear vowels. There was also a 50% increase in air flow, and the vowels were longer. Fischer-Jørgensen suggests that in addition to the widening of the rear part of the glottis, there is also an increase in the activity of the expiratory muscles which keeps the sub-glottal pressure up and accounts for the absence of any loss in intensity.

In the light of the above, let us turn to Punjabi, which has developed tones in the environment of Indo-Aryan *bh (but not *b). The following summary is culled from accounts by Bahl (1957, 1958) and Chatterji (1940: 113-114), where the tones are impressionistically described, and not instrumentally documented. However, they are in substantial agreement on the facts. Syllable initial *bh gives a low rising/falling tone, and the initial consonant has de-aspirated (and de-voiced word-initially). Using the system scale developed for Chinese, it appears to have contour of the form , say 24. An example would be *bhukh > púkkh. When a stressed syllable ends in a voiced aspirate, the tone is high falling (written \check{v}), and the consonant de-aspirates, for example: *baddha > b\`dda. When a

syllable begins and ends in voiced aspirates, the tone is described as 'two tones side by side' by Chatterji: *bhābhī > pá:bi. Chatterji also notes that loss of h produces a tone, and the tone is accompanied by a quantity of breath (i.e., breathy vowel) 'which may be a characteristic feature of the tone'.

We see, then, that languages unrelated to Chinese have apparently developed lower tone in the context of voiced aspirates and the resulting breathy vowels. This is of course different from the registers splitting in an already tonal language like Chinese; we shall see that both effects of voiced aspiration can be seen in the complex splits and mergers that make up the development of Mandarin tones.

(c) The Tai Case

Li (1977a) reconstructs the following four-way contrast for Tai:

*p *ph *b *ʔb

Of these, only *b gives rise to the 'lower' register tones, whereas *ʔb behaves like the glottal stop and gives rise to 'upper' register tones. He says (p. 26):

"It is generally believed that a voiced consonant initial had the tendency to depress the tone, so that the two tone series were distinguished by pitch level."

However, the generalization that voicing causes the split in tone series cannot be maintained given Li's own reconstruction, since the pre-glottalized consonants, which have upper register reflexes, were also voiced, as Li notes on p. 68:

"The reconstruction of *ʔb is based on the fact that all Tai dialects show some voicing, although the glottal feature may have been lost."

Using the Halle-Stevens laryngeal features, we see then that only voiced stops which are [- constricted glottis] caused register split in Tai, and voicing is a necessary but not a sufficient condition for register split. It may in fact be possible to go still further and suggest that only [+ spread glottis] voiced stops caused the split (exactly as in the Chinese dialects), but to maintain this, detailed investigation by someone familiar with the Tai dialects is needed. Li includes some evidence that seems to make such a reconstruction plausible. Firstly, the reflex of Li's *b is aspirated in many modern dialects, as one would expect if it were in fact *b^h (and Li's *ʔb were in fact simple *b). Li argues that this aspiration is a secondary development. Secondly, the development of the voiced liquid *r into h in many dialects causes Li to suggest that it was in fact heavily aspirated (p. 142). We might therefore suppose that it was in fact *r^h; however, no equivalent evidence exists for the other sonorants. All this is speculation; what is clear is that voicing alone was not responsible for the register split: either aspiration, or its opposite, glottalization, must be invoked as well. Voicing was a necessary but not a sufficient condition for register split in Tai just as in Chinese.⁷

(d) Other Cases

Two African languages provide examples of similar processes taking place productively in the phonology. The first case comes from Feʔfeʔ, a Cameronian Bantu language. The data come from a talk

given by Larry Hyman (MIT 1978). Roughly speaking, stops and some other consonants become aspirated before high vowels. This occurs with voiced as well as voiceless stops, and when a voiced stop is aspirated in this way before high tone, it has the effect of depressing the start of the tone, giving a rising contour. Voicing, then, is necessary but not sufficient for lowering initial pitch.

The second case is more familiar. Data are from Lanham (1963). Xhosa initial consonants can be divided into two classes -- depressor and non-depressor consonants. Depressor consonants are so named because they lower the beginning of the following tone, changing high to rising, and falling to a convex rising-falling tone. The depressor class are all voiced, but some voiced consonants fall into the non-depressor group. Among the obstruents, voiced implosives are non-depressors, while plain voiced stops are depressors. Among the nasals, plain voiced nasals are non-depressors, while voiced aspirated nasals are depressors.

Depressors	b	m̥			ɸ
<hr/>					
Non-depressors	ɓ	m	p	ph	h

The general statement that the voiced consonants are depressor consonants is thus untrue: voicing is a necessary but not a sufficient condition for the depressor effect.

Lastly, I should like briefly to discuss Tibetan. Data are from Kjellin (1977). On the surface Tibetan has two tones, a high and a rising. Both tones do not occur with all possible initial consonants, and Kjellin in fact argues that the tones are fully predictable from

the underlying consonants. Whether or not this is right is not relevant to what follows. The initial consonants occur as follows:

<u>High Tone</u>				<u>Rising Tone</u>		
p	ph	m		b	ph	m
?	h			h		

I am completely unfamiliar with the history of Tibetan, and therefore do not intend to make any claims about the origins of the two tones.⁸ I would simply like to say that it is interesting to note that the above picture is exactly what one would expect to find if the rising tone came from proto-voiced aspirates *bh̥, which had either de-voiced to ph, or de-aspirated to b, m, but never both (accounting for the absence of rising tone syllables starting with p). Further, there would have to have been a contrast between aspirated and unaspirated voiced nasals, now visible only in different tonal reflexes. The obvious competing analysis would propose voicing as the source of the split. In that case there must have been voiceless nasals (and in fact there still are), and presumably both aspirated *bh̥ > ph, and *b > b, with de-voicing limited to the aspirates. This is also perfectly plausible.

To summarize this section, we have seen a number of examples from a variety of language families in which voicing is a necessary but not a sufficient condition for two different processes: register split, and the lowering of the beginning of pitch contour. These two effects have a logical explanation; the first phenomenon, register split, affects the pitch of the entire syllable, and we have already

seen that breathy vowels in Gujarati are of significantly lower pitch than their clear equivalents. We might therefore suggest that register split is caused only indirectly by the laryngeal features of the initial consonant itself, via the laryngeal quality of the following vowel. The change in pitch contour, on the other hand, is a direct effect of the consonant on the beginning of the following vowel. We shall see both at work in the next section.

3.3.3 Mandarin and Cantonese Tonal Splits

Cantonese provides an almost perfect picture of a tonal system which has split across the board into two registers. At the beginning of the seventh century A.D., it is known that Middle Chinese (from which most modern Chinese language families are taken to be descended) had four 'tonal' categories. Opinions differ as to whether or not these categories, or Sheng, were truly tonal, or whether they were still distinguished by different types of syllable closure (glottal, aspirated, plain, stopped) as they were at a yet earlier stage. This will be discussed at some length below. All that need concern us here is that at some point the four categories split into eight. All syllables which began with voiceless obstruents in Middle Chinese (henceforth MC) have Yin (or Upper) register reflexes; all syllables which began with voiced obstruents have Yang (or Lower) register reflexes. What was the nature of these voiced obstruents? Cantonese no longer has voiced obstruents, and the voiceless reflexes of the earlier voiced series are aspirated or unaspirated depending on the tonal category. In those Chinese dialects (such as Shanghai) which retain voiced obstruents they are usually aspirated (and accompanied

by breathy vowels), and for this reason, among others, most authorities reconstruct MC voiced stops as having been aspirated. Some linguists, including F-K Li (1977b) have pointed out that consonant systems in which the voiced series is aspirated but not in contrast with an unaspirated voiced series are highly marked, and have therefore reconstructed MC as having unaspirated voiced stops. Such an assumption leaves unexplained the reflexes of this series in modern dialects, and we will therefore accept the reconstruction of aspirated voiced obstruents.

The results in Cantonese are as shown overleaf. Note that the secondary splits in Yin Ping and Yin Ru are later developments and irrelevant here. The Yin tones occur with both aspirated and unaspirated initials, but the Yang tones have only the type shown on the chart. Mandarin has undergone a slightly different set of changes, and has subsequently lost the final consonants of the Ru category, which has been re-distributed as shown.

In this section I will try to account for two things:

(1) The presence or absence of aspiration in the Yang tones in modern Mandarin and Cantonese.

(2) Which tones split and which ones merged.

I will not account directly for the actual contours of the modern dialects, since these are the result of later developments of great complexity leading to enormous variety (see example chart in section 2.2.1). The output of my tone rules are rather kinds of proto-Mandarin and proto-Cantonese, respectively.

The argument that follows will take the following form. The best available evidence suggests that Middle Chinese was already truly

The following two charts illustrate the situation in Mandarin and Cantonese.

Mandarin

MC Tonal Categories

	Level	Oblique		
	Ping	Shang	Qu	Ru
YIN	┘ 55	┘ 21	┘ 41	┘ 21
*p *ph				(=Yin Shang)
YANG	┘ 35		p	┘ 35
*bh	ph			(=Yang Ping) p

(re-distributed)

Cantonese

MC Tonal Categories

	Level	Oblique		
	Ping	Shang	Qu	Ru
YIN	┘ 53/55	┘ 35	┘ 44	┘ 5 (ṽ)
*p *ph				┘ 4 (ṽ)
YANG	┘ 21	┘ 24	┘ 33	┘ 3
*bh	ph	ph	p	p

Note that Ping, Shang and Qu were sonorant (vowel or nasal) final syllables but Ru syllables ended in voiceless stops (and still do in Cantonese).

tonal. Therefore, although I accept the basic proposal of Pulleyblank (1978) to the effect that breathy vowels induced by initial voiced aspirates developed into low register, I do not accept that the presence or absence of initial aspiration is the result of assimilatory and dissimilatory processes between initial consonant and final closure. Rather, I propose that aspiration was lost during the operation of a rule which changed tonal contour under the influence of that aspiration.

Let us now consider the evidence in more detail.

(a) Middle Chinese Tonal Categories

In order to explain their subsequent development, we obviously need to consider the Middle Chinese tonal categories themselves. Some brief remarks were made in section 2.2, and these are repeated and expanded here. Their forms and contours in the modern dialects are so diverse that we can as yet learn little from that source. Instead, we return to contemporary sources as studied by Mei (1970) and Pulleyblank (1978). Both writers agree on the following points:

- (i) the Ru category had a final stop (or fricative, see later)
- (ii) 10th century poets separated out Ping 'Level' from all others 'Oblique'
- (iii) Ping was used to transcribe Sanskrit long vowels
- (iv) Shang was used to transcribe short vowels

The major points of disagreement are:

- (i) was Qu longish or shortish?
- (ii) was Shang high level or rising?
- (iii) was the Level/Oblique distinction one of Low/High or Long/Short?⁹

Mei's paper was the earliest serious attempt to reconstruct MC tones. In addition to the use of the tones to transcribe Sanskrit vowel length distinctions, he refers to a Japanese description of Chinese tones of the period, and the tradition of bombai, Sanskrit psalmody transliterated into Chinese and brought to Japan, with rules for chanting. Pulleyblank uses these and related sources, but discounts the bombai evidence for reasons that are not clear to me. Mei considers that the use of Qu to transcribe long vowels (albeit with the addition of a rider 'longish') precludes the possibility of the Level/Oblique contrast being based on length. He further analyzes the Shang tone as high; and thus takes the Level/Oblique distinction to have been Low/High. Pulleyblank, on the other hand, notes that Qu was sometimes used at an early stage in MC to transcribe short vowels, and proposes that it had a final voiced \underline{h} , while Shang had a final glottal stop. Ping is then the only category ending in a clear vowel, and as such might be expected to be longer than the Oblique categories. The two reconstructions are diagrammed below:

	LEVEL	OBLIQUE		
	Ping	Shang	Qu	Ru
Mei	low long level	high short level	?high longish rising	?high short (VC) ?
Pulleyblank	V	V?	Vh	V \hat{c}

Obviously, a decision between these two proposals depends largely on whose analysis of the original documents one accepts, and this demands a level of scholarship I do not possess. Pulleyblank's

reconstruction is relatively noncontroversial for a very much earlier period (say, up to the 4th century A.D.) but no descriptions for the 6th century onwards mention final ʔ or ɦ, although in other respects they are very detailed. Pulleyblank's analysis (see below) is also apparently not transferable to the modern Kwangsi dialects, where the register split is happening now, with essentially similar effects on aspiration, yet no such final segments are present.

(b) Pulleyblank's Analysis

In many ways Pulleyblank's assumption about the MC tonal categories allows for a very elegant analysis of their subsequent development. He postulates two processes, represented below:

- I Glottal Assimilation bhVʔ → bhVɦ
- II Dissimilation of Voiced Aspiration bhVɦ → bVɦ

and assumes that the tone split was triggered by breathy vowels.

(I will return later to the question of whether the vowels in these rules should be represented as breathy or not.)

These two processes work in the first three tones as shown to give the Mandarin facts:

	<u>Ping</u>	<u>Shang</u>	<u>Qu</u>
	bhV	bhVʔ	bhVɦ
I	-	bhVɦ	-
II	-	bVɦ	bVɦ

Since initial voiced aspiration gives a breathy vowel, the Ping tone will split into two registers under the influence of this. Shang has

already separated under the influence of Rule I, and merged with Qu in the Yang register. Qu does not split, although in fact Pulleyblank explicitly says (p. 21) that the breathy vowel induced by the initial voiced aspiration is still present, so a split might be expected. We will return to this later. At a later stage, by the way, there is general devoicing of obstruents and the final closures of Shang and Qu disappear.

The Ru category requires a further assumption. If these are reconstructed as is customary, with a final plain unaspirated stop, then the loss of aspiration cannot be explained, since the context for rules I and II will not be met. So Pulleyblank reconstructs final glottalised stops (or fricatives) which then undergo generalised versions of rules I and II:

		bhVʔ
I		bhVʰ
II		bVʰ

Notice that the effect of Rule I is again to split the registers, as with the Shang category.

To summarise in chart form:

LEVEL	OBLIQUE		
Ping	Shang	Qu	Qu
phV	phV?	phVh	phVʔ
pV	pV?	pVh	pVʔ
bhV	bhV?	bhVh	bhVʔ
I -	bhVh	== -	bhV ^h ʔ
II bhV	bVh	== bVh	bV ^h ʔ

In Cantonese, the main difference is the failure of Rule I to operate in the case of the Shang category. Rule II can therefore not apply, and aspiration remains. The Shang tone splits into two registers under the influence of the initial voiced aspiration in the same way as the Ping does. Notice that the development of the Qu category is identical for both dialects, and yet it splits in Cantonese but not in Mandarin.

LEVEL	OBLIQUE		
Ping	Shang	Qu	Ru
phV	phV?	phVh	phVʔ
pV	pV?	pVh	pVʔ
bhV	bhV?	bhVh	bhVʔ
I -	-	-	bhV ^h ʔ
II -	-	bVh	bV ^h ʔ

(c) Problems with Pulleyblank's Analysis

1. MC syllables could end in vowels, nasals, or stops.

Pulleyblank must set up a three way contrast in each case:

V	V?	Vh
Vn	V ^h	V ^h
Vt	V ^t	V ^t

and so on for the labials and velars. Most of these are known to be possible consonant types -- the only one of which I know no examples is the voiced aspirated nasal finally (initially, of course, they occur in Chinese dialects). However, a three way contrast in the nasals is unknown, I think, and was unremarked at the time, despite the usual detail of contemporary descriptions.

2. There is some doubt as to whether syllables of the form $b\dot{h}V$ and $b\dot{h}V^h$, both with breathy vowels, are really distinct. Certainly Ken Stevens would be surprised to find them contrasting in a language. Since they crucially distinguish the Ping and Qu categories in Pulleyblank's analysis, they must have been distinct for MC speakers.

3. The split in registers in Qu in Cantonese but not Mandarin is a real problem. Morris Halle suggested to me that "Grassman's Law" (Rule II) might operate differently in the two dialects, simultaneously de-voicing in Mandarin but not in Cantonese. The voicing would then split the registers. This is the only possible solution, and it invalidates the claim that both voicing and aspiration are necessary for a register split in Chinese. Notice that if we instead suggest that Mandarin Qu does not split because the final \dot{h} caused Yin as well as Yang vowels to be breathy, we still cannot

explain why Cantonese Qu does split unless we also assume the variation in Rule II suggested by Halle.

4. The failure of Rule I (and therefore Rule II) to apply in Cantonese Shang is a problem, since the rule is present in the grammar, and operates in Ru. We must therefore suppose that Cantonese had a restricted version of Rule II operating only on final obstruents.

Informally:

$$bhV\acute{t} \longrightarrow bhV\acute{h}$$

5. I mentioned earlier that Mandarin subsequently lost the final stops in the Ru category, and re-distributed the words among the other tonal categories. This re-distribution was regular, and must be accounted for (but see (f) below).

Yang Ru	$*bhV\acute{t} \longrightarrow pV\acute{h}$	goes to	Yang Ping	$*bhV \longrightarrow phV$
Nasal Ru	$*mV\acute{t}$	goes to	Yin Qu	$*pV\acute{h}$
Yin Ru	$*pV\acute{t}$	goes to	Yin Shang	$*pV?$

The last merger is not surprising -- once the stop was lost, the two syllables would be indistinguishable (either both having final glottal stop, or similar tones developed from that). The Yang Ru, however, might by the same argument be expected to merge with the Yang Qu, which had a final h. Instead they merge with Yang Ping, and they do so overwhelmingly across dialects (Dell 1977). Pulleyblank gives the following explanation. First he sets up two features for Early Mandarin's six tonal categories, breathiness (which might perhaps be considered low tone instead) and length:

	Yin Ping	Yang Ping	Yin Shang	Qu	Yin Ru	Yang Ru
breathy	-	+	-	+	-	+
long	+	+	-	-	-	-

Now both Ying Shang and Yin Ru are $\begin{bmatrix} -\text{breathy} \\ -\text{long} \end{bmatrix}$, and they merge. But we would still expect Yang Ru, presumably also $[-\text{long}]$, to merge with Qu. So Pulleyblank says "Evidently it (Yang Ru) must have lengthened at some time during its evolution" (p. 27). He even advances a possible explanation, suggesting that the final voiced aspirate stops or fricatives resulting from the application of Rule I might have been relatively longer than the glottalised consonantal closures of the other Ru tone words.

Now the nasals merged with Qu, which is $[+\text{breathy}]$; they must therefore have had voiced aspiration originally, in which case they would have been subject to Rule I in its most general form, and the final closure would have become aspirated. However, they obviously failed to undergo the lengthening process just discussed, since they merge with the $[-\text{long}]$ Qu. There is no obvious reason why this should be the case.

(d) A Tonal Hypothesis

We have seen that there are certain problems, although none of them insurmountable, with Pulleyblank's approach. Let us now investigate the consequences of assuming, with Mei, that Middle Chinese had true phonemic tones, which interacted directly with aspiration in some way. (Note that Mei does not tackle the aspiration facts; he is solely concerned with reconstructing the tones on the basis

of contemporary evidence.)

I mentioned earlier one obvious advantage of this approach -- the Kwangsi dialects have phonemic tones, and yet some of these processes are apparently going on at this moment.

If we accept Mei's reconstruction, we might represent MC tones as a sequence of tones, each of which can be high or low, $[\pm H]$, as follows:¹⁰

Ping	Shang	Qu	Ru
$[-H -H]$	$[+H +H]$	$[-H +H]$	$[+H -H]$

The level/oblique distinction is then one of the presence or absence of high tone at some point in the syllable. Those features which will be taken to be suprasegmental in Chinese include tone, aspiration, and perhaps voicing; note that these are all laryngeal features, so their interaction is not unlikely.

In addition to the above assumption about MC tone contours, I shall propose two rather natural processes. The idea is that breathy vowels, which are known to be lower in pitch than clear vowels, give rise to register splits -- that is, the whole vowel lowers in tone.

On the other hand initial voiced aspiration makes its pitch lowering effect felt only at the beginning of the vowel, as is usual with initial consonants, and so it alters the contour of the tone by depressing the beginning, rather than changing the register. At the same time, the aspiration is lost.¹¹

First, let us formalise the rules. Register splitting is a process of spreading the feature [+ Spread Glottis] (+SG) onto the vowel, giving a breathy vowel. To put it another way, [+SG] is becoming autosegmental.

	Ping	Shang	Qu	Ru
	[-H -H]	[+H +H]	[-H +H]	[+H -H]
(2)	{ -	[-H +H]	[-H +H]	[-H -H]
	{ -	b	b	b
(1)	-U / \	-	-	-
	-H -H			

In Cantonese the order is reversed, and rule (2) is of a slightly different form, as noted above.

	Ping	Shang	Qu	Ru
	[-H -H]	[+H +H]	[-H +H]	[+H -H]
(1)	-U / \	-U / \	-U / \	-U / \
	-H -H	+H +H	-H +H	+H -H
(2)	{ -	-	-U / \	-U / \
	{ -	-	-H +H	-H -H
			b	b

(e) Problems Solved and Problems Posed

1. Problems one and two -- problems of the phonetic plausibility of Pulleyblank's account -- disappear under this analysis.

2. The third problem -- why does the Qu category split in Cantonese and not in Mandarin -- is here seen to be the result of a difference in rule ordering, since in Mandarin the contour changing rule bleeds the register splitting rule.

3. The fourth problem -- the failure of glottal assimilation to apply to the Shang category, even though it does apply to the very similar Ru -- is here shown to be the result of a more restricted

contour changing rule in Cantonese than in Mandarin. In Cantonese the rule applies only to contoured tones (i.e., [α H] [$-\alpha$ H]), whereas in Mandarin it applies to all 'oblique' tones, that is everything except low level [-H -H] tones. This is captured in rule (2) by adding the option of [+H] to term 3:
$$[\alpha H] \begin{cases} [-\alpha H] \\ [+H] \end{cases}$$

4. Next, we come to the mergers of the Ru tones in Mandarin. Pulleyblank's analysis correctly predicts the mergers of Yin Ru words, and nasal-initial Ru words, but he is forced to postulate secondary lengthening to account for the Yang Ru merger with Yang Ping. This is exactly what is predicted by our analysis, since both are [-H -H]. Conversely, however, Yin Ru and nasal initial Ru words are now a problem. Yin Ru is [+H -H], and merges with [+H +H] Shang. Nasal-initial Ru are [-H -H] (assuming they were aspirated), and merge with [-H +H] Qu. In both cases, then, we must postulate a secondary raising of the end of the tone, and this raising is conditioned by the initial consonant, since obstruent-initial Yang Ru ([-H -H]) words do not undergo it. Not only is this messy, but the little phonetic evidence I know of which bears on this would predict a lowering rather than a raising of the end of the tone. Facts from Lea (1973) on English, and Stahlke (1971) and a phonological process in Ewe (cited in Hyman (1976)) suggest that final obstruents lower the pitch of a preceding vowel. On the other hand final ? is supposed historically (see Haudricourt 1954) to have caused a rising tone on the preceding syllable by raising the end of the tone, and since the stopped Ru tones usually passed through a ?-final stage before total loss this

could be the source for the final raising.

We shall see in (f) that there is some evidence that the behaviour of Yin Ru and nasal-initial Ru words should be accounted for quite separately as a later development. If this is the case, we may take the correct prediction of Yang Ru behaviour as an advantage of our analysis, and the complications with the rest of the Ru category as an entirely separate problem.

5. In general, then, the tonal analysis presented here accounts for the same facts as does Pulleyblank's analysis -- the distribution of aspiration in Yang tone words, the merger of Yang Shang and Yang Qu in Mandarin, the split in the Ping tone in Mandarin; in addition, it accounts for some facts which he fails to account for -- the split in Qu in Cantonese but not Mandarin, the loss of aspiration in Shang in Mandarin but not Cantonese, the merger of Yang Ru and Yang Ping. Further, a tonal analysis fits contemporary descriptions of Middle Chinese tones very closely. The Register/Tone distinction is central in the analysis: no other tonal approach known to me can properly explain these complex facts.

(f) A Note on the Tone of *ʔ and *∅ Words

With very few exceptions, *ʔ words have Yin tones, and *∅ have Yang tones. The fact that [+ Constricted Glottis] initial words are Yin is hardly surprising given our hypothesis, but the Yang tone of *∅ initial words needs some explanation. Such words are found only in the so-called "third division," which means they in fact began with a front glide i or y.¹² We have already noted in a footnote that one of Kwangsi dialects has breathy vowels after glides in low register words;

we therefore suggest that MC glides, like all other voiced segments, could be aspirated, and that in fact the so-called * \emptyset initial words began with such breathy glides. These contrasted with so-called * ʔ initial words, which may better be described as beginning with glottalised glides. Those words which had no glide, but a glottal stop, were not in contrast, and the glottal stop can be considered a surface phenomenon. The glides would have the following features (English glides are included for comparison):

	Spread Glottis	Constricted Glottis	Stiff Vocal Cords	Slack Vocal Cords
English w,y	-	-	-	-
English W,Y	+	-	+	-
MC (uh)ih,yh	+	-	-	- (?)
MC(u?)i?,y?	-	+	+	-

Note that Pulleyblank must assume something similar for his analysis.

Hsueh (1975) notes a problem for any approach which tries to treat all the tonal splits as a single phenomenon. He points out that, whereas in Ping, Shang and Qu tone words * ʔ initial words are consistently Yin, in the Ru tones in Mandarin they group with the nasals, liquids and * \emptyset , subsequently merging with Qu. He therefore suggests that * ʔ must have been lost (i.e., merged with * \emptyset) after the changes in the Ping, Shang and Qu tones, but before the changes in the Ru. We may therefore take the Ru tone mergers to be a late phenomenon, and it need not concern us (for the purposes of this section) that the Yin Ru and nasal-initial Ru mergers remain problematical. I have investigated the behaviour of such words in Cantonese, and facts are

quite different. *ʔ initial words show up as Yin in all tones, including Ru. I suspect that a secondary development of some kind has produced the Mandarin situation.

A note about sonorant initial words in general seems needed. Such words are found in both Yin and Yang tones, but the majority are Yang. It is apparently also the case that a word may be Yin in one dialect and Yang in another, which would mean that it was aspirated in the second dialect only. Far from being a problem, this is exactly what we would expect given the modern dialects of Shanghai and Kwangsi. Sherard (1972) points out that although nasals may be unaspirated, the aspirated variety are very much more common. The Kwangsi dialects show that when voiced aspiration is lost it is lost in the sonorants first, and only later in the obstruents. Obviously then MC had already started to lose aspiration in the sonorants, although the majority were still aspirated at the time these changes took place. Why this should be the case, however, remains an open question.

Professor Pulleyblank has pointed out to me that the sonorants did not participate in the Shang-Qu tone shift in any dialect. This suggests that the nasals (in Shang at least) may not have been truly aspirated, but rather had simple breathy vowels. They would then still be subject to register split, but not to contour change. Since contour change is responsible for the Shang-Qu shift, they would never participate in this. Note, however, that Rongxian, one of the Kwangsi dialects discussed by Tsuji, apparently has aspirated nasals in Shang as well as the other categories, and the Wu dialects lack even breathiness in Shang, so the picture is very unclear.

To summarize again, the development of Mandarin and Cantonese provides a complex illustration of the interaction between voiced aspiration and tone. As in the other cases we have seen, voicing is a necessary but not a sufficient condition for register split. The involvement of aspiration is particularly clear in this example because the interaction was mutual, with aspiration itself being lost under certain tonal conditions (that is, when contour change operates). It has proved possible to explain the development from MC to a hypothetical early Mandarin and Cantonese, but we have made no attempt to continue our exploration right up to the modern dialects; that task must wait for a future researcher. It is worth taking a look at the magnitude of the problem, however.

In the chart below we give tones for all the Cantonese dialects given in McCoy (1966).

	Yin Ping	Yang Ping	Yin Shang	Yang Shang	Yin Qu	Yang Qu	Yin Ru	Yang Ru
Rule output	33	11	55	22	35	24	53	11
Szeyap dialects (incl. Stand- ard Toisan)	33	11	55	21	33	32	5 3	32
Standard Cantonese	53	21	35	23	33	22	5 3	2
Popei	55	13	33	45	53	31	43 22	31 55
Zhongshan	55	31	13	13	33	33	55	33
Kau Sai	53	21	35	22	44	22	5 3	2
Tan Shui Hau	33	35	13	13	31	33	3 2	4
Yeung Kong	33	331	31	31	13	131	13 31	131

It is clear that extracting general trends is exceedingly difficult, but it is interesting to compare the tones of the Szeyap dialects with the output of the rules proposed here (see first line in chart). In the first four cases the match is exceedingly good, and in the last two it is quite plausible. Only the absence of a rise in the Qu tones is a problem, and the rise is missing in all the dialects.

3.4 The Physiology of Pitch Control

Any system of distinctive features must of course be grounded in an understanding of their articulatory and acoustic correlates, but for pitch control this has proved a remarkably difficult task. The only serious attempt is that of Halle and Stevens (1971), which attempted to unify the laryngeal phenomena of voicing, aspiration, glottalization and tone by means of a single set of features, and thereby explain some of the observed connections between consonantal features and the pitch of the following vowel. The features of consonants and vowels are given below:

	1	2	3	4	5	6	7	8	9
Obstruents	b_1	b	p	p_k	b^h	p^h	ʃ	?b	?p
Glides	w,y				h	h,W,Y			?,?w,?y
Vowels	V	\check{V}	\acute{V}	voiceless V	breathy V			Creaky V	Glottalized V
Spread Glottis	-	-	-	+	+	+	-	-	-
Constricted Glottis	-	-	-	-	-	-	+	+	+
Stiff Vocal Cords	-	-	+	-	-	+	-	-	+
Slack Vocal Cords	-	+	-	-	+	-	-	+	-

The foundation of the Halle and Stevens system was the assumption that the laryngeal muscles could control two things independently: the size of the glottal opening, and the tension on the vocal cords. A single laryngeal gesture could have several effects; for example, slackening the vocal cords caused voicing on consonants, but also lowered the pitch of vowels. This explained why both historically and synchronically voiced consonants were known to have a pitch lowering effect on the adjacent vowel (although the effect is much smaller, perhaps close to or below the level of perceptability, in tone languages. See Hombert, Ohala and Ewan (1979)).

The major problem with this system is the fact that it can only account for three pitch levels (columns one to three) instead of the four (or five) that are needed in languages. Furthermore, none of these can be combined with any voice quality other than ordinary clear voice, since the same features that account for pitch account for changes in voice quality. Contour tones or pitch distinctions within, for example, breathy vowels can therefore not be described in this system. For a more detailed criticism of other problems see Anderson (1978: 161-166).

Another attempt was made by Brown (1965), in which he set out to explain why a single consonant type could apparently raise the pitch of the vowel in one set of dialects and lower it in another (the so-called Voice High and Voiced Low Tai dialects). He suggested that the answer lay in different timing of the release of the adductor and abductor muscles of the larynx, but the details of his proposal are unclear and there is no evidence that I am aware of to support

this proposal.

One problem is that in many cases not only is the mechanism for pitch control unknown, but the phonetic facts are also unclear. For example, the effect of (voiceless) aspiration on the pitch of the following vowel is quite erratic, and some researchers have found a raising effect while others have found a lowering effect, often within the same language.

Hombert (1978: 87-90) gives a useful review of the problems in this area, and after additional experimentation concludes that the reason for the contradictory findings is that:

"the parameters controlling the rate of vibration of the vocal folds are not significantly different at the onset of a vowel following a voiceless aspirated stop versus a voiceless unaspirated stop."

In the case of voiced aspiration (or breathy voiced consonants, as some researchers feel that the term aspiration is misleading in connection with voiced stops) the phonetic effects are clear: they lower the pitch of the following vowel. But the mechanism is still open to discussion. Hombert speculates on the contributory factors:

"Although the rate of airflow is high upon the release of breathy voiced consonants, the vocal cords are not closely adducted, and thus the Bernoulli force should be weak. In addition, during breathy voice there is a less forceful contraction of the laryngeal adductor muscles..., many of which not only act to bring the vocal cords together but also are known to participate in F_0 regulation..."

In Hombert, Ohala and Ewan (1979: 48) he adds that:

"The sub-glottal pressure is also markedly lower upon their (i.e., the vocal cords) release (because of the high rate of air-flow), and this by itself would lead to a somewhat lower F_0 ."

Ken Stevens (personal communication) has speculated that the answer might lie in a somewhat different direction. When the glottis is fully opened, the vocal folds have further to travel to the almost touching position; each cycle might thus take longer, giving a greater period of vibration and a lower frequency.

By now it should be clear that phonetic motivation for any system of tone features must await a clearer understanding of the mechanisms of production (and perception). Ultimately the link to the phonetics must be made, however, and without it any feature system is partly speculative.

NOTES TO CHAPTER THREE

¹ Such a language is Moulmein Pho, a Karen dialect (Jones 1961: 69) which has only H and L tones on morphemes and no contours. Japanese may also be such a case.

² Also Yoruba, in which apparently only L pulls down H, not M.

³ LM sequences are excluded in the language for other reasons, so it is not possible to tell if L pulls down M or not.

⁴ Vietnamese as expounded by Haudricourt (1954) is the major example of this type of which I am aware. I do not know whether proto-Vietnamese is considered to have had *b^h or *ʔb; if so, it would be most interesting to know their subsequent development.

⁵ I should like to thank Professors Li Fang-Kuei, Mei Tsu-Lin, and E. G. Pulleyblank for reading an earlier version of this section and making a number of extremely helpful comments and criticisms. All errors of course are my own.

⁶ Tsuji also notes that glide initial syllables in low tones may or may not be breathy voiced, although he also maintains they are never 'murmured' (i.e., aspirated). This of course means that breathy vowels and low tone cannot be fully predicted from the presence of murmur, although he claims they can. He does not offer instrumental evidence to support the statement that glides may have breathy vowels, but not voiced aspiration, and it seems possible that they may in fact be

slightly aspirated.

⁷ Brown (1965, 1975) showed that in addition to the Tai dialects which had developed lower register after *bʰ there were also a large number of dialects like Lao which had done exactly the reverse, with low register after *pʰ and high register after *bʰ. Various explanations have been advanced for these dialects, including the suggestion that originally they too had voiced initials giving low register, but underwent some kind of flip-flop that reversed the correspondence. A more interesting proposal has been made by John Grima (as reported in Strecker 1979: 57ff.). He notes that the Mon-Khmer languages developed not tone but rather a voice quality difference as follows: voiceless initials produced creaky voiced syllables and voiced initials produced clear or ordinary voiced syllables. Since creaky voice (like breathy voice) is associated with lowered pitch, this would be a plausible intermediate stage for the Lao dialects where voiceless initials produced low tone, whereas the Tai dialects with low tone after proto-voiced initials evolved from a stage where those voiced initials caused breathiness on the following syllable.

	C difference	→	Register difference	→	Tone difference
voiced-high:	paa 24	→	<u>paa</u> 24	→	paa 13
	baa 24	→	paa 24	→	paa 24
voiced-low:	paa 24	→	paa 24	→	paa 24
	baa 24	→	<u>paa</u> 24	→	paa 13

The likely involvement of creaky voice in tonal development is supported by Rose (1979). Rose points out that in Zhenhai, a Wu dialect, the historically voiced initials have given rise to breathy high vowels but creaky non-high vowels. In both cases F_0 is lower during the breathy/creaky part of the vowel than on a clear vowel.

⁸ This is apparently a different idiolect of Lhasa Tibetan from the one more commonly described in the literature, which has only a contrast between p and ph instead of the contrast between b and ph reported by Kjellin (Ballard, personal communication).

⁹ Mei Tsu-Lin (personal communication) has pointed out to me that the level/oblique distinction dates from at least the sixth, and perhaps the fifth century; however, the philological evidence on which phonetic reconstruction is based is eight century at earliest, and the tones may well have changed in the interim. However, since the distinction remained active in use, it seems to me that we may take it to represent a real and continuing difference.

¹⁰ [+ H] should be understood as shorthand for some bundle of laryngeal features.

¹¹ Professor Pulleyblank has pointed out to me (personal communication) that there are dialects like Su-chou which have retained voicing (and aspiration?) but underwent the Shang-Qu shift. It would appear that in such dialects contour change (which caused tone shift, see below) and loss of voiced aspiration were independent processes. There is no reason why dialects should not differ in this respect, since the

connection is plausible but not necessary. If this is right, such dialects might differ in that Qu has a Yin/Yang split, unlike Mandarin. I do not know if this is in fact the case.

¹² Professor Pulleyblank has also pointed out (personal communication) that it is not in fact true that all third division words with the 'zero' initial had a medial glide. It seems, then, that we must accept the existence of some underlyingly breathy vowels in addition to those conditioned by the initial voiced aspirate.

4. THE AUTOSEGMENTAL NATURE OF REGISTER

In chapter two a number of arguments were advanced to show that tone is autosegmental in a variety of Chinese languages. In this chapter we will use the same arguments to show that each feature of tone must be treated as a separate autosegmental tier. In many cases the same data that provided the basis for the arguments in chapter two can now be given a more refined analysis that allows us to separate out the features of Register and Tone as independent prosodies. In other cases new data will be introduced. In order to avoid repeating all the arguments in detail the reader will often be referred to the relevant section in chapter two.

4.1 Partial Morphemes

If Register and Tone are independent tiers, both of which are also independent of the syllabic level, then morphemes should be able to lack one or more of the levels. In addition to morphemes fully specified on all levels (the unmarked case) we should therefore expect to find morphemes lacking Register, Tone or the syllabic level, and morphemes consisting of only Register, Tone or the syllabic level. Examples of all except one of the above can be found in Chinese languages, and the absence of that one -- morphemes lacking just the Register component -- is probably not significant but just the result of chance. I would expect such morphemes to exist, and eventually they will probably be found. We have mentioned that the unmarked case is for morphemes to be fully specified on all levels; the most highly marked case, on the other hand, is for a morpheme to have a

specification for Register or Tone but not both. Intermediate between the two extremes are morphemes specified for both Register and Tone (but not syllables) or for syllables (but not Register or Tone). These include the most common cases of floating tones, and toneless morphemes. To put it another way, the laryngeal features of Register and Tone are usually both specified, or neither may be, but it is much more highly marked to have a partial laryngeal specification only.

4.1.1 Morphemes Lacking One Level Only

We have already drawn attention to the accidental gap in our data that includes no morphemes lacking only register. For an example of a morpheme that lacks only Tone we turn to Mandarin neutral tone (see also 1.2.1).

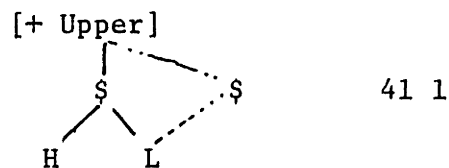
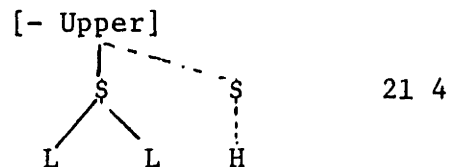
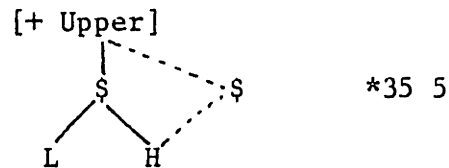
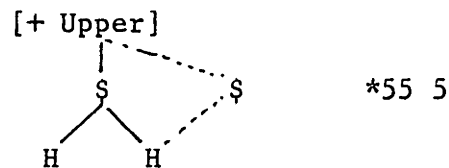
The four tones of Mandarin will be represented as follows:

First:	55	[+ Upper] $\begin{array}{c} \diagup \quad \diagdown \\ H \quad H \end{array}$
Second:	35	[+ Upper] $\begin{array}{c} \diagup \quad \diagdown \\ L \quad H \end{array}$
Third:	21	[- Upper] $\begin{array}{c} \diagup \quad \diagdown \\ L \quad L \end{array}$
Fourth:	41	[+ Upper] $\begin{array}{c} \diagup \quad \diagdown \\ H \quad L \end{array}$

(The choice of HH and LL instead of H, L is based on the stress facts and also the formulation of the sandhi rules given in 5.1.)

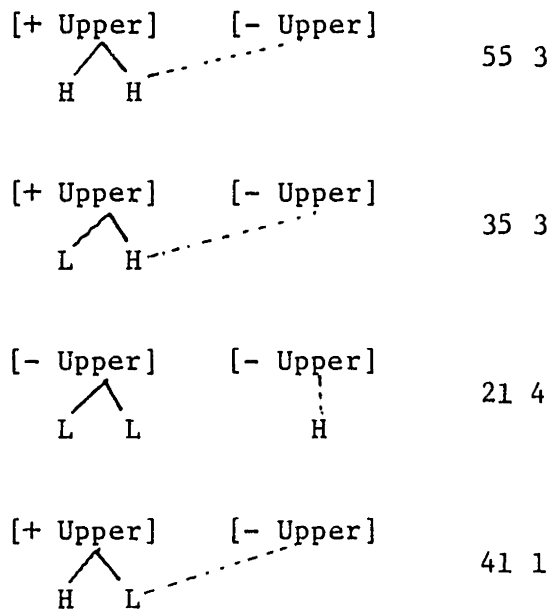
There is a special rule that inserts a H tone after the third tone when no other tone follows (i.e., pre-pausally, or before a neutral tone). Now recall that there are various morphemes that never show up with a tone of their own, but are always in neutral tone, and that their pitch is therefore predictable from the tone of the preceding syllable. It is however not the case that their pitch can be explained by straightforward spreading of both Tone and Register from the preceding syllable.

If that were the explanation, we would have the following representations (after insertion of the H after the third tone, and attachment by the WFC):



This gives the right results in the case of the third and fourth tones (assuming that the lowness of the end-point of the fourth tone is a phonetic detail. For an argument that this is so, see section 5.1.), but it gives the wrong results in the first two cases: the right output has a mid-level tone on the second syllable.

What we suggest is that these morphemes lack only the Tonal tier; the Register tier is specified [- Upper]. This immediately gives the right results:



We therefore have lexical entries like the following:

[- Upper]
le 'PERFECTIVE ASPECT suffix'

that lack only a Tonal tier. As a result the overall level of such morphemes is always low, but it is maximally low after a syllable ending in a L Tone, and mid-level after a H Tone, since the melody spreads from the preceding syllable.

The next kind of morpheme to be discussed lacks only a syllabic level: in other words it consists of Register and Tone only. Examples of this sort are quite common, and include the floating tones discussed so far. In section 1.1.1 we showed that the changed tone in Cantonese can be best understood as a high floating tone that is suffixed to the base morpheme. The result is that any word made up of a morpheme beginning on a high level plus this suffix ends up consistently high level, but any word whose first morpheme begins on a non-high level ends up rising to high:

$$\begin{array}{ccc}
 \left. \begin{array}{l} 53 \\ 55/5 \end{array} \right\} & \longrightarrow & 55/5 \\
 & & \\
 & & \left. \begin{array}{l} 44/4 \\ 33/3 \\ 22 \\ 35 \\ 24 \end{array} \right\} \longrightarrow 35
 \end{array}$$

It is easy to see that this morpheme must be specified for both Register and Tone, since its pitch is independent of the Register or Tone of the preceding syllable but always remains [+ Upper, H], even after a [22] tone, which is presumably [- Upper, L]. We conclude that this is an example of a morpheme lacking only the syllabic component, and that there are lexical entries of the following kind:

[+ Upper]

'DIMINUTIVE suffix'

H

A similar argument can be made for the morpheme used in Amoy triple

reduplication (see 1.1.3), which must also be [+ Upper, H] with no syllabic specification.

It is not the case that such morphemes are always [+ Upper, H], and in section 5.2 we will see that Shanghai has a kind of boundary marker that is [- Upper, L] with no syllabic element. Interestingly I know of no cases that come in between these two, which may be significant.

4.1.2 Morphemes Lacking Two Levels

All three of the logical possibilities are found: morphemes consisting only of the syllabic tier, only Register, or only Tone. Examples of the first kind have already been encountered in section 1.4.2 in Shanghai, although not discussed explicitly. We said that in Shanghai words only the tone of the initial syllable is relevant, and all others are deleted. However, in many cases it is not clear that it makes sense to postulate a tone as ever having been present. Suppose that a morpheme never occurs in isolation or word-initially, and the speaker does not know how it is written. There are many such morphemes, and they never surface with an identifiable tone of their own but rather take it from the preceding syllable. The initial consonant may narrow down the possible tones that the syllable could have, but it will still leave some ambiguity, and in the case of sonorant initial syllables the ambiguity can be total. This is because sonorant initial syllables can occur with any tone, the only constraint being that in tone three, the low rising tone, breathy voice or murmured release is always present. However, in non-initial position not only is the tone lost but so is breathy voicing, so the second syllable in the examples below could have any of the three tones that occur in Shanghai on unstopped

syllables (two clear and one breathy):

快 悔	ɔ̃ lɔ	'regret'
○ ○	ghəq nəŋ	'thus'
一(点)	yq n̄ɛ	'a little'
○ ○	hhi la	'they'

Sherard (1971: 131) says that native speakers have a tendency to associate sonorants with breathy voice (statistically more frequent) and to guess on that basis that they have the breathy tone underlyingly; he therefore writes the above examples with h on the second syllable even though it is phonetically absent. (This runs counter to the common claim that in such circumstances speakers will pick the surface form, which is non-breathy here.) The fact remains that any choice of underlying Tone or Register in these cases is arbitrary, just as it is arbitrary to assign a Tone (although not a Register, which is always [+ Upper] for unvoiced obstruents) to the second syllable in these cases:

自(家)	zhɪ ka	'by oneself'
○ ○	lhɔ pɔ	'a madame, mamasan'

since they could either be high falling tone one, or high level tone two.

We conclude that the lexical entries for sonorant-initial morphemes of this type for at least some speakers lack both Register and Tone, and are of the following form:

$$\left[\begin{array}{c} [- \text{Upper}] \\ \text{hi} \\ \text{L} \quad \text{H} \end{array} \right] + \left[\begin{array}{c} \text{la} \end{array} \right] \quad \text{'they'}$$

(The breathy voice on the first syllable is predictable from [- Upper], and therefore is not included in the lexical representation.)

For examples of morphemes consisting only of Register and only of Tone we turn to some new facts from Mandarin intonation. The following facts are taken from Chao (1968); a more detailed account of Mandarin intonation that differs in many details from Chao's can be found in Rumjancev (1972).

One of the intonation patterns described by Chao provides a fairly clear example of a morpheme consisting solely of [- Upper] Register. This is the 'conclusion' intonation (p. 39-41) that characterizes the second half of a bi-partite utterance of the following kind:

Nǐ xìng Wáng, wǒ xìng Lù 'Your name is Wang, my name is Lu'

你 姓 王 , 我 姓 陸

Wǒ xìng Lù, nǐ xìng Wáng 'My Name is Lu, yours is Wang'

我 姓 陸 , 你 姓 王

Lù is a fourth tone word, the high falling 41. Wáng is a second tone word, the high rising 35. Chao says of the fourth tone:

"the falling intonation will make it start lower and squeeze it narrower"

and of the second tone:

"...Wang will be pitched lower as a whole, but still with a second tone contour..."

Unfortunately he is not explicit about the effect on the other two tones, but it seems that they are realized at a lower level than usual, as would be expected if they were all [- Upper] Register. Note that the second

and third tones would still remain distinct for reasons of timing, since they will become:

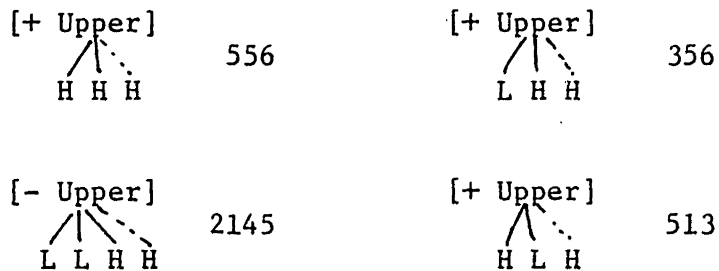


respectively in pre-pausal position, and we might therefore expect the rise on the third tone syllable to start much later than the rise on the second tone syllable. This is indeed the case, as Rumjancev makes clear. For this reason I will claim that although [LL] and [L] merge phonetically, [LLH] and [LH] do not: timing is crucial. Note that this analysis of conclusion intonation assumes that the intonation morpheme, [- Upper], actually substitutes for the underlying Register of the final syllable. The alternative would be to assume that the phonological Register is unchanged, and it is rather the realization rules that determine the phonetic pitch of the Register that are affected. For an analysis of English intonation along these lines see Pierrehumbert (1979).

There are also two intonational-type morphemes consisting solely of Tone discussed by Chao in the chapter on sentence particles (pp. 812-4). One is a rising ending that expresses incredulity, impatience, or peremptory command. The other is a falling ending that expresses a superior or condescending attitude that Chao compares to the low rising English intonation in effect. The forms of these endings are as follows:

	<u>Rising</u>		<u>Falling</u>	
			<u>On single syllable</u>	<u>On syllable + neutral</u>
First:	55	556	552	55 21
Second:	35	356	352	35 31
Third:	214	2145	2141	21 41
Fourth:	51	513	5121/5122	51 121/52 21

Now consider how these facts are to be explained. The rising intonation results in a rise of more than one digit in only one case: when the final syllable is a fourth tone syllable. The level it rises to is then a mid-level, whereas after the first two tones it stays up very high (and may go even higher). Suppose that this morpheme consists of only a H Tone, with no syllabic elements and no Register. Then we have:

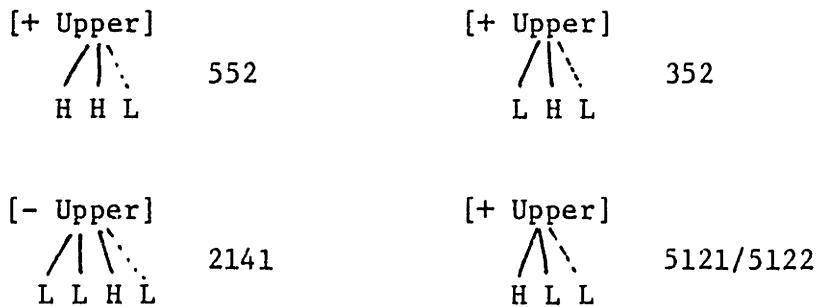


This correctly predicts essentially the right forms, except for the final pitch in the case of the fourth tone, where we might expect 535 or 515. The fact that it does not rise all the way is undoubtedly linked to the fact that fourth tone syllables fall lower than one might expect for an Upper Register tone (i.e., to 1 instead of the expected 3); presumably the extreme contortions of pitch implicit in a representation like 515 are subject to some kind of performance

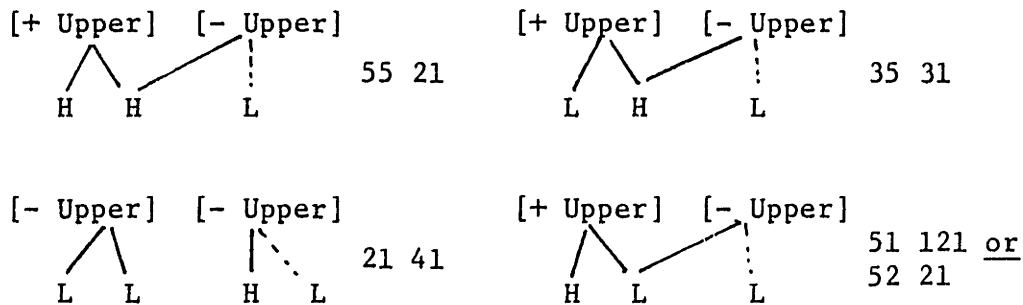
constraint that results in only partial realization of the rise. Pace this problem, then, the assumption that the rising ending is a sentence particle consisting of a Tone and no syllables or Register correctly accounts for the facts. The lexical entry would have the form:

$$\left[\begin{array}{c} \\ H \end{array} \right] \quad \text{'particle expressing incredulity'}$$

Now consider the falling ending. This embodies the other logical possibility for a single floating Tone: a L Tone with no Register or syllabic level. The result of attaching this particle will be as follows in the case of a single syllable:



If the final syllable is neutral toned, we have:



Note that the association lines as drawn here assume that association of the intonational morpheme happens after insertion of the H after

the third tone, and after spreading onto the neutral toned syllables. This is the most natural assumption, since undoubtedly the attachment of a sentence particle like these endings will be in a later cycle. The above outputs are again essentially right, producing a lower Register falling tone on the final syllable in all except the last case, when it is more-or-less low level. The variation between 21/31/41 in this instance is taken to be non-significant.

We therefore have another lexical entry consisting of tone only:

$\left[\begin{array}{c} \\ L \\ \end{array} \right]$	'sentence particle expressing condescension'
---	--

4.2 Partial Deletion

If Register and Tone constitute independent levels it should be possible not only to delete segmentals and leave tone behind, but also to delete Register or Tone but not necessarily both. Such cases do indeed exist, and we will discuss one example of each. Further examples can be found in the full analyses of several languages in chapter five, but will not be given here because they depend on fairly intricate arguments which would be out of place here.

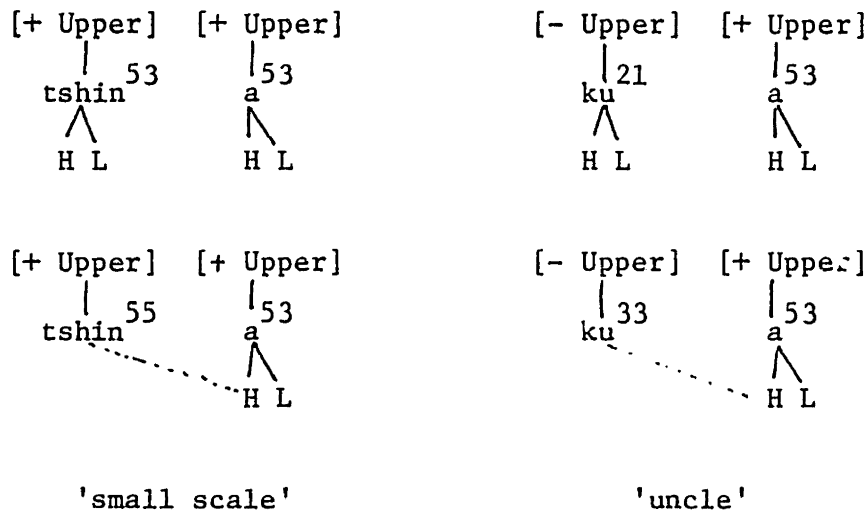
4.2.1 Deletion On One Level Only

For the first example let us consider some facts from Amoy that have already been discussed in section 1.4.3, and which provide an example of a process which deletes Tone but leaves Register and the syllabic level untouched. These are the tonal variants found before the suffix a⁵³:

	2	3	4	8	7	5	1
Pre-pausal	53	21	21?	54?	33	13	55
Contextual	55	53	54?	21?	21	33	
Before a ⁵³	55		5?	3?	33		

There are basically only two variants before this suffix, one high level and one mid-level. Which one occurs apparently depends on the overall level of the contextual tone: it is certainly independent of its shape (level or falling), and its end-point (53 and 33 have different outputs). This suggests that the Register of the tone is unchanged by the addition of the suffix, and the observed effects must be due to changes in the melody.

It turns out that if we make the simple assumption that Tone is deleted before this suffix the observed outputs are immediately accounted for. This is because if Tone is deleted the WFC will require that the Tone of the suffix spread backwards onto the preceding morpheme. Now the suffix has a high-falling tone, which must be [+ Upper, HL]. The Tone that spreads will therefore be H, and this will interact with the underlying Register of the morpheme to produce level 5 in the case of a [+ Upper] tone, and level 3 in the case of a [- Upper] tone:



It should be clear that there is thus a rule of Tone deletion in Amoy that leaves both segmentals and Register untouched. Such a rule is only possible if Tone and Register are separate autosegmental levels.

For a rule that deletes only Register we must turn to a language that has not been discussed so far. This is Fuzhou, a Northern Min language, and it exhibits an exceedingly complex set of tone sandhi that changes the tone on a morpheme in different ways depending on what tone follows. The data here are drawn from several sources (which disagree in many ways; see section 5.4 for a discussion) including T'ao (1930), Chao (1933) HYFYCH (1964), HYFYGY (1960), Wang (1969), Lan (1953).

When the first syllable has glottal closure (shown by an underscore in the table below) the glottal closure is lost in some cases but not in others, as follows:

	<u>Second syllable:</u>						
	44	52	<u>4</u>	22	12	<u>13</u>	242
<u>First syllable</u>							
<u>4</u>	44			22			
<u>13</u>	22			35	<u>4</u>		

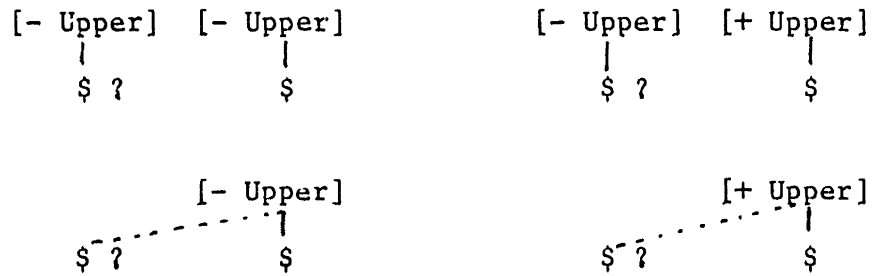
We shall see in section 4.4 that tones 44, 52, 4 and 22 must be considered [+ Upper] Register because of their effect on vowel height. It is thus the case that the final glottal is always lost on a [+ Upper] syllable /4/, but on a [- Upper] syllable /13/ only when the following syllable is + Upper. We might therefore write:

$$? \longrightarrow \emptyset / \left. \begin{array}{l} \text{—————} \\ [+ \text{ Upper}] \\ \text{—————} [+ \text{ Upper}] \end{array} \right\}$$

These two contexts are remarkably similar in that both refer to [+ Upper], and we do not have to look far for a way to collapse them so that only the first environment is needed. For this to be the case it is necessary that underlyingly [- Upper] Register stopped syllables become [+ Upper] Register before [+ Upper] Register syllables only. But this is exactly what would happen if such syllables had lost their own Register and were therefore subject to the WFC which would spread the Register of the following syllable backwards. To see this, let us formulate the relevant rule:

$$[- \text{ Upper}] \longrightarrow \emptyset / \text{—————} [+ \text{ closed glottis}]$$

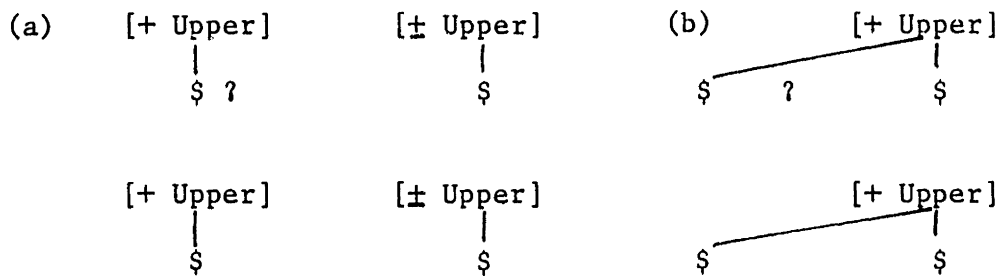
Now observe the effect of this rule and the subsequent application of the WFC:



In the left-hand column there has been no effective change, but in the right-hand column the output has merged with underlying [+ Upper] Register syllables. It is now only necessary to state a single environment for glottal deletion:

$$? \longrightarrow \emptyset / \text{-----} \\ [+ \text{ Upper}]$$

This rule will then apply to the derived [+ Upper] cases in the right-hand output above, and to underlying [+ Upper] Register syllables, but not to the left-hand output above:

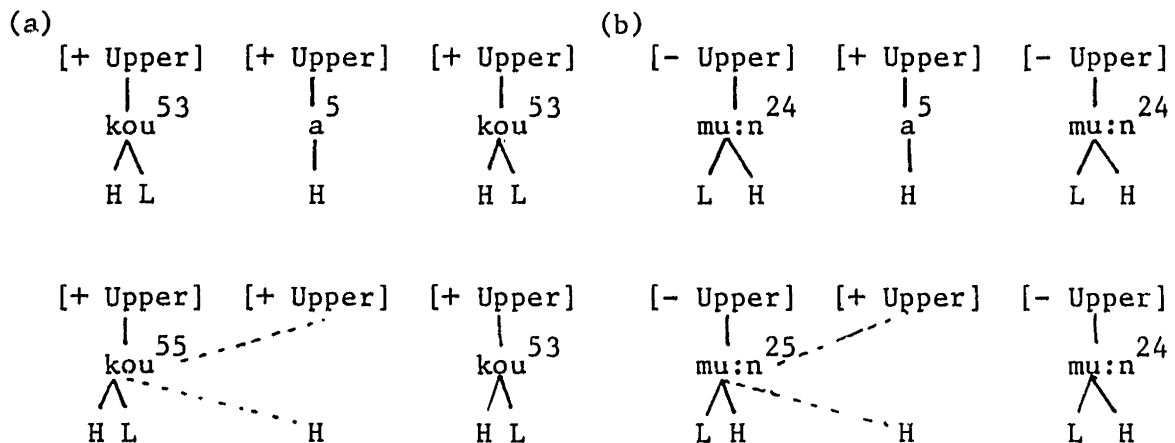


Note that both glottalization and pitch control are laryngeal phenomena, and their interaction in a rule like this is therefore unsurprising.

A rule of [- Upper] Register deletion therefore allows for a

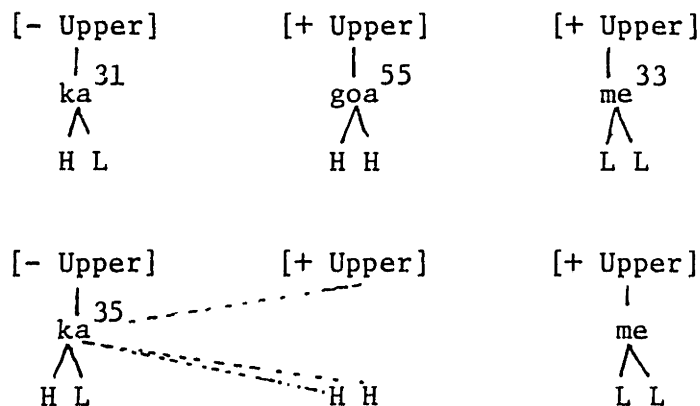
simpler statement of glottal stop deletion than would otherwise be possible, and such a rule is only formulable if Register and Tone are independent levels. (For a full explanation of the eventual output tones see section 5.4.)

The last type of deletion on a single level is deletion which takes only the syllabic level, leaving both Register and Tone behind. This is probably the unmarked case, and includes the cases of segmental deletion discussed in section 1.3. To recapitulate, in Cantonese there are certain morphemes which can be deleted under certain circumstances, but they leave a tonal effect behind. The deleted morphemes all have either high tone, or a high rising tone, and this reattaches leftwards resulting in a changed tone on the preceding morpheme. This changed tone always ends on a maximally high level (level 5) no matter what the underlying Register or Tone of the preceding morpheme. We must therefore conclude that what is left behind is both [+ Upper] Register and H Tone (or LH Tones in the case of high rising morphemes being deleted) and only the syllabic level is deleted. The resulting derivations are exemplified below:



With the addition of a very natural rule simplifying sequences of more than two Tones on a single syllable by deleting the medial Tones, the correct outputs are automatically obtained.

The facts of Amoy contraction (see 1.3.2) suggest a similar analysis. For example, contraction can result in the coalescence of tones from different Registers and of different melodies, and the output forms are exactly as predicted under the assumption that only the segmentals are deleted while both Register and Tone remain:



(again, given a simplification rule).

Having shown that any one of the three levels may be deleted independently of the others, it now remains to discuss whether more than one may be deleted simultaneously.

4.2.2 Deletion On More Than One Level

The prediction of autosegmental theory is that the simplest rules will be those that act within a single level; however, it is not the case that rules cannot make reference to other levels, or even act on more than one level simultaneously, but just that such cases will be more highly marked. One entirely reasonable proposal might be

that in the standard case purely phonological rules will apply on one level, whereas morphological rules might apply on all levels. In particular, while phonological rules might delete elements from a single tier, morphological rules might delete entire morphemes. This is the case in Sanskrit, for example, but it is possible to see immediately that this hypothesis cannot be maintained as an absolute restriction by looking at the Cantonese changed tone discussed in the preceding section, and in 1.3.1. The deleted morphemes are undoubtedly deleted by a morphological rule rather than a phonological rule, and yet it is not the case that the entire morpheme is deleted: rather only the syllabic level disappears. Similarly Amoy contraction is morphologically rather than phonologically conditioned, and yet the tones stay behind. A second consequence of the hypothesis that divides rules up into two types -- those that delete one level only, and those that delete whole morphemes -- is that there should be no rules that delete on two out of three levels. Below we will discuss one such case, and there are others in chapter five. Notice however that the two levels which are deleted are always Register and Tone, which suggests that there is a closer link between these two levels than between either of them and the syllabic layer: that is, one would perhaps not expect to find rules which delete Tone and syllabics, but not Register, or vice-versa. No such cases have come to my attention, and I suspect that they are at least highly marked.

The example we will discuss here is Shanghai deletion (see also 1.4.2). As we have already pointed out, word-internally in Shanghai only the first syllable's tone is relevant, and all others are

deleted. Not only is the melody irrelevant, but the Register, which is manifested not only in pitch but also in breathy voice in Shanghai, is also deleted. As a result morphemes which have breathy voice in isolation or in initial position in a word lose their breathiness in word-internal or word-final position and merge with syllables that never show breathiness. For example, as first syllables in compounds hú 火 is [+ Upper] Register, clear voice, while hhu 扶 is [- Upper] Register, breathy voice. But the homonymous morphemes hú 貨 and hhu 榻 are both spoken with a normal clear voice as the second element in a word:

火車	hú cho	'train'	扶梯	hhu thì	'ladder'
退貨	thé hú	'contraband'	闖禍	chóng hhu	→ chong hu '(auto) accident'

It is therefore necessary to assume that both Register and Tone are deleted, and only syllabics are left behind.¹

The last type of example one might expect to find would be a rule that deleted the whole of a morpheme. Such a rule would only be detectable if the morpheme had first had some effect on the adjacent morphemes so that a trace of its erstwhile presence was left. The rather limited interaction between morphemes in most Chinese languages makes such cases hard to find, but one possible example has come to my attention. There is a Wu dialect called Hǎiyán Tōngyuán which has frequent deletion of the numeral 'one' in a wide variety of morphological contexts. The facts which follow are from Hú (1959) and were brought to my attention by Matthew Chen. The tones of the constituent morphemes of a word in Tongyuan depend on the number of

syllables in that word. For example, words beginning in a rising [25] toned morpheme will have the following tones on subsequent syllables:

Two syllables: 25 + 42

Three syllables: 25 + 42 + 31

For example, the verb $khə^{25}$ 看 'to look' can be reduplicated for tentative aspect either once or twice, and the results are as follows:

$$\begin{array}{l} khə^{25} khə^{25} \longrightarrow khə^{25} khə^{42} \\ khə^{25} khə^{25} khə^{25} \longrightarrow khə^{25} khə^{42} khə^{31} \end{array}$$

However, there is a third possible tone pattern just in the case of reduplicated action verbs. The meaning of this pattern is equivalent to the meaning of the Mandarin reduplicated verb with numeral 'one' in the middle, but in Tongyuan no numeral ever surfaces; instead there is a new tone pattern:

$$\backslash \quad khə^{25} khə^{25} \longrightarrow khə^{25} khə^{31} \quad \text{'take a look'}$$

The simplest way to account for these facts, as Hu suggests (p. 374), is to assume that "the medial 'one' must be deleted." This will account for the tone of the final syllable, which is what would be expected for the last syllable of a tri-syllabic form:

$$khə^{25} i?^5 khə^{25} \longrightarrow khə^{25} i?^{42} khə^{31} \longrightarrow khə^{25} khə^{31}$$

Note that not only the syllable level, but also the tonal levels must have been deleted, because they do not surface after re-attachment to an adjacent syllable.

There are many other examples of the same type in other environments, and with other tones. For example, the normal tone patterns after a short [2]-toned syllable are given below:

Two syllables: 2 + 213

Three syllables: 2 + 213 + 434

But if a verb is followed by a measure word the numeral one is omitted but the tone of the last syllable shows that the form must originally have been tri-syllabic:

$$\begin{array}{l} \text{tshə}^2 + \text{i}^5 + \text{khei}^{434} \longrightarrow \text{tshə}^2 + \text{i}^{213} + \text{khei}^{434} \\ \longrightarrow \text{tshə}^2 + \text{khei}^{434} \quad \text{'eat a mouthful' } \text{吃 (-) 口} \end{array}$$

The last example has an initial syllable with a [242] tone, and normally that results in the following patterns:

Two syllables: 242 + 213 (when the second syllable is /434/)

Three syllables: 242 + 434 + 31

But after deletion of an intermediate 'one' we get:

$$\text{zu}^{242} + \text{ti}^31 \quad \text{'sit down for a moment' } \text{坐 (-) 点}$$

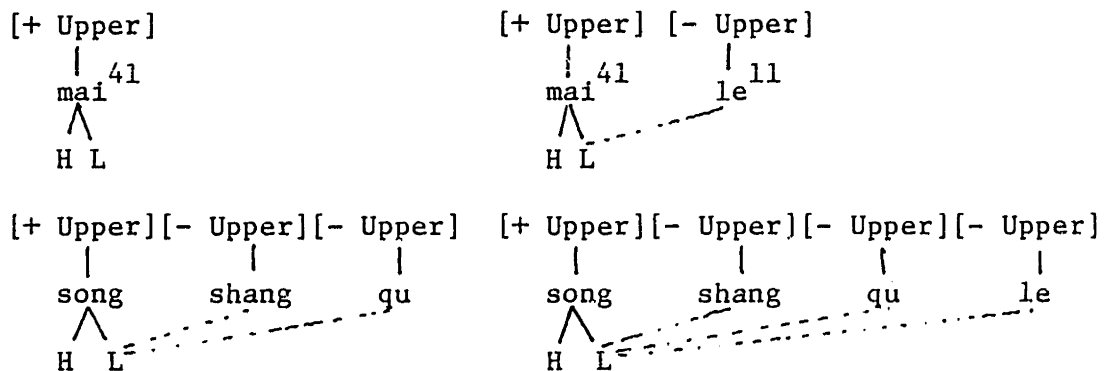
with the last syllable in the tone usually reserved for the last syllable of a tri-syllabic form.

I have not attempted a full analysis of the tone sandhi of Haiyan Tongyuan, and it may turn out that the deletion involves more than one rule. If it is the result of a single rule, however, as seems likely, that rule must delete the entire morpheme-syllabic

level, Tone and Register. Since the rule is morphologically conditioned this is not at all surprising, but rather constitutes the unmarked case.

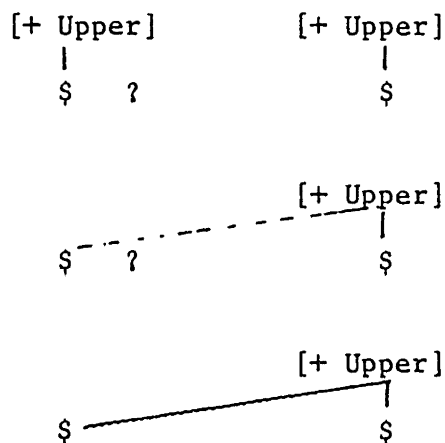
4.3 Spreading

A related claim of autosegmental phonology is that the melody of a form will remain stable under segmental changes. If both Register and Tone are autosegmental, then each should remain independently stable, and spread over any unassociated syllabic elements. It is possible to see this process if one looks at cases of morphemes that lack one level in the lexicon, or as the result of a deletion rule. For example, the spreading of Tone happens in the neutral tone in Mandarin, so that the Tonal melody stays stable in all the following cases:



(see also 1.5 for other cases.)

The spreading of Register happens in Fuzhou, and creates the environment for glottal stop deletion (see 4.2.1):



If longer forms were to exist, the prediction is clear: Register will continue to spread leftwards onto any Registerless syllables.

In other words, stability of melody is just as applicable to the separate autosegmental levels of Tone and Register as it is to a single tonal tier as represented in Goldsmith. In many cases both spread at once, for example in Shanghai words, where both Tone and Register spread as far right as necessary.

4.4 The Domain of Register and Tone

I have claimed that underlyingly Tone may occur in sequences, but Register does not change over the morpheme. There is one possible consequence of this claim that has not been mentioned so far. Consider a syllable associated with two Tones and one Register; each Tone will be realized over roughly half the syllable, whereas the Register will be realized over the entire syllable. If then there is any interaction between either tonal feature and some other part of the syllable we might expect that each Tone could only interact with its own end of the syllable, whereas Register might interact with the entire syllable.

Let us make this more concrete. In many Chinese languages stopped syllables can only bear a limited number of tones, and these are most often level tones on the surface. Sometimes, however, there is reason to believe that they are underlyingly contour tones because they behave as such in tone sandhi. This is true in Fuzhou (see 5.4) for example, where it can be shown that the high stopped 4 tone is actually HL, and a late rule raises the end of the tone. This rule must make reference to the syllabic structure if the stopped tones end in a velar k rather than a glottal stop, as they do in some cases. The rule must therefore be stated roughly as follows:

$$L \longrightarrow H / \begin{array}{c} \text{Vk} \\ | \quad \diagdown \\ \hline \quad \quad * \end{array}$$

I should like to claim that rules of this kind will always apply to the relevant end of the syllable, and that it is not necessary to write the rule as above, but rather that it can be simplified to:

$$L \longrightarrow H / \begin{array}{c} \text{Vk} \\ | \\ \hline \end{array}$$

This will then always be interpreted as meaning the final tone of the syllable, and never as:

$$* L \longrightarrow H / \begin{array}{c} \text{V k} \\ | \quad \diagdown \\ \hline \quad \quad \text{T} \end{array}$$

Such rules should, I maintain, be impossible.

Now consider the difference for Register, which always covers the

entire syllable. If Register were to affect the syllabic level or vice-versa it should affect the entire domain; rather than interaction with the initial or coda, then, we might expect interaction with the nucleus. Certainly interaction with the coda (like the above) is not known to me, and there is one particularly clear case of interaction between the Register and the nucleus which we will now present.

Fuzhou's chief claim to fame among Chinese languages is that it shows a set of alternations in vowel quality under tonal conditioning. These rather interesting facts are laid out below (the data are from T-C Wang (1969)):

Tones:	44	52	<u>4</u>	22	35	12	<u>13</u>	242
Vowels:			i				ɛi	
			y				oey	
			u				ou	
			ɛi				ai	
			oey				əy	
			ou				ɔu	

These correspondences hold not only for underlying forms, but are preserved under tone sandhi (in one direction, in any case. See below). If a morpheme bearing one of the tones in the second column changes into one of the tones in the first column, the vowel changes also:

$tɕi^{242} huoŋ^{44} \longrightarrow ti^{44} uoŋ^{44}$
 $toey^{12} tsɔuk^{13} \longrightarrow ty^{52} zɔuk^{13}$
 $hou^{242} ny^{22} \longrightarrow hu^{52} ny^{22}$
 $kaiŋ^{242} tyoŋ^{22} \longrightarrow kɛiŋ^{52} nyoŋ^{22}$

(The consonantal changes are irrelevant here.)

Sources differ on whether tones ever change in the opposite direction: an underlying /52/ becomes either [22] or [12], depending on the authority (see also T'ao (1930), Chao (1933), HYFYCH, HYFYFY, Lan (1953) Egerod (1956)). If it becomes [22], both input and output are always accompanied by higher vowel alternants so no change would be expected. If it becomes [12], which normally takes lower vowel alternants, we might expect the vowels to lower, but in fact they do not:

$ti^{52} tɔ^{12} \longrightarrow ti^{12} lɔ^{12}$ NOT $tɛi^{12} lɔ^{12}$
 $ku^{22} san^{44} \longrightarrow ku^{12} lan^{44}$ NOT $kou^{12} lan^{44}$

Let us therefore write a raising rule as follows:

$$V \longrightarrow \left[\begin{array}{l} - \text{low} \\ - \alpha \text{ high} \end{array} \right] / \text{---} \quad \text{in tones } 44, 52, \underline{4}, 22, 35$$

[αlow]

The question now is how to define the environment, since the list is clearly unsatisfactory. Maddieson (1976) has claimed that

'a correlation of pitch height with vowel height is not possible'

(p. 198) because it is not the case that all the tones which trigger

the rule are higher than all those that do not. In particular, the [22] tone triggers the rule but the [242] tone does not, and neither does the [13] tone (which Maddieson takes to be [24], following Norman (1973)). But given the feature system developed in this thesis there is no problem whatsoever. Suppose that the context for the rule is simply [+ Upper] Register:

$$(1) \quad \begin{array}{c} v \\ [\text{low}] \end{array} \longrightarrow \left[\begin{array}{c} - \text{ low} \\ - \text{ high} \end{array} \right] / \frac{\quad}{[+ \text{ Upper}]}^2$$

This requires that tones 44, 4, 52, 35 and 22 are all [+ Upper], an entirely uncontroversial claim for all but the last. Notice however that within the system [22] is the mid-level tone, and this is presumably more significant than the fact that it is usually written [22] rather than the more common [33] for mid-level tones. There is therefore no reason not to assume that it is [+ Upper, L] and indeed we will see in 5.4 that the complex tone sandhi of Fuzhou are easily explained under this assumption.

The remaining tones can then be represented as [- Upper], and in 5.4 we will justify the remaining features of these tones.

Having shown that there is no reason why the context for the rule cannot be stated in very simple (in fact minimally simple, using a single feature) tonal terms, it is necessary for completeness to dispose of Maddieson's second argument against such a rule. He suggests that although there is some phonetic motivation for a link between vowel height and high pitch that might explain how such a rule came into being, it turns out on closer inspection that it is

not only vowel height that is involved, but also rounding assimilation and backness assimilation, for which a phonetic explanation of the tonal context would be hard to find.

Maddieson bases his argument on data which differ in certain respects from that given above. The two sets of facts are compared below:

Maddieson	Wang		Maddieson	Wang
ei	ɛi		i	i
øi	œy		y	y
ou	ou	→	u	u
ai	ai		ei	ɛi
oi	øy		øy	œy
au	ɔu		ou	ou

Maddieson then points out that in addition to the raising process two other changes can be observed: assimilation of rounding when /oi/ → oy (→ øy) and assimilation of fronting in the first element, as when (/oi/ →) oy → øy. Consider the rounding assimilation first. Maddieson's data involve three cases of this:

/oi/ → oy (→ øy) øi → øy (→ yy → y)
 /au/ → ɔu (→ ou by raising)

The first case is dubious; most other writers (Egerod, T'ai, Wang, HYFYCH) give front rounded vowels for both variants with the glide held constant as either i or y. The second case only arises for

Egerod, who has no /y/ in the coda. The third case is probably a true case, since HYFYCH, T'ao and Egerod all give [au], not [ɔu] (Wang has [ɔu]). Notice that for all these writers the language has both [a] and [ɔ] elsewhere. This rounding must however be by a different rule from that responsible for the earlier cases, since the rounding goes in the opposite direction. Further, it is entirely independent of the sandhi changes, since we have non-alternating word-final [au] (even Wang considers this to be [au], not [ɔu]). If, then, we set up underlying /ɔ/ we will correctly account for those dialects with an alternation [ɔu] ~ [ou], and those which show [au] ~ [ou] probably have a rule unrounding low vowels before [u]. It remains only to say that the dialect described by Egerod also has a rule unrounding front glides after vowels: $\phi y \rightarrow [\phi i]$. I can see no good reason for associating either of these facts with the tonal context: they are independent aspects of the phonology. Now consider the frontness adjustment proposed by Maddieson. This is limited to the single case cited above, and as already stated most authorities give no such alternation, but rather $\Theta y \rightarrow \phi y$, where Θ is a low frontish rounded vowel. Even if, following Chen and Norman (1965), there is a change in backness, there still seems no good reason for considering it part of the same process. It is at least as reasonable to assume that the sandhi involves raising only, and that in some cases subsidiary processes are responsible for different surface forms. So I would propose underlying / Θy /, going to [ϕy] in sandhi, but with a possible backing rule for low rounded vowels. This will give / ϕy /, the form given by Wang. Front glide unrounding in some dialects will give

/ɔi/. Two facts in Maddieson's data remain somewhat mysterious (although susceptible to ad hoc treatment). Firstly, he actually has [oi], not [ɔi]. Secondly, front glide unrounding is rather idiosyncratic for this dialect, apparently. The above discussion will account for all the dialects in the sources referred to in this section, so I shall take it as demonstrated that the small amount of dialectal variation is independent of the tonally conditioned raising process, and that this can be described by the rule in (1).

Before leaving this topic I should point out that it is not necessary to assume that only Register can interact with the syllable nucleus, but rather that this is a result of the lack of sequences in Register. If a language had tones that consisted of a single Register and a single Tone, either of these could interact with the syllable nucleus, and this has certainly happened historically. For example, Cantonese originally had two-way pitch distinction of stopped syllables and this distinction (as we have argued in section 3.3.3) was one of [+ Upper] Register. However, at some later stage in most of the Yue dialects, including Cantonese [+ Upper] Register tones underwent a further split conditioned by the quality of the vowel. The conditioning feature has been argued to be [+ Tense] (see Yue 1976), and the result was that tense vowels developed lower tones than lax vowels. This is still the situation in modern Cantonese, although in most of the other modern dialects the tense/lax distinction has since been lost or obscured. In one dialect, Bo-bai, the split took place in the [- Upper] Register also (perhaps under the influence of long/short vowels), but in all the others it was limited

to the [+ Upper] Register. If the situation before the split was that these syllables were associated with only one Tone and only one Register, the rule can be simply stated as:

$$H \longrightarrow L / \begin{array}{c} [+ \text{ Upper}] \\ | \\ [+ \text{ tense}] \\ | \\ \hline \end{array}$$

This will then result in [+ High] tones in the Upper Register.

We would not expect to find a situation where [+ tense] vowels triggered a contour change by changing only one of a sequence of Tones, or at least such cases should be highly marked. Yue (1976) discusses one such possibility in the Lien-chou dialect where apparently [- high] tense vowels caused the tone to become [+ rising] (from low level) (see p. 54), and to the extent that her analysis is correct we must weaken our claims somewhat to allow rules like:

$$\phi \longrightarrow H / \begin{array}{c} [- \text{ high}] \\ \swarrow \quad \searrow \\ L \quad \hline \end{array}$$

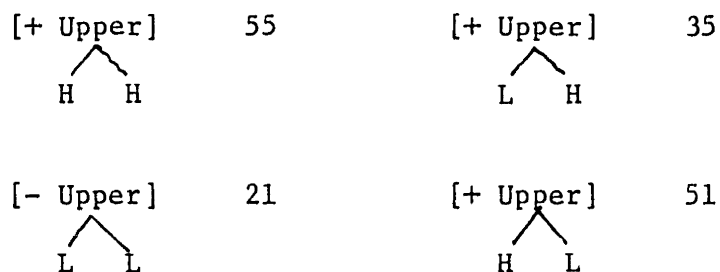
NOTES TO CHAPTER FOUR

¹ Unfortunately Sherard (1971) does not give examples with the same morpheme in initial and final position, but he is quite explicit that neutralization does take place.

² The rule requires some segmental restriction, since the low vowels only raise in finals ending in a consonant ([ŋ] or [k]), although HYFYG disagrees.

5. SUMMARIES OF THE TONAL PHONOLOGY OF VARIOUS LANGUAGES5.1 Mandarin

We suggested in section 2.1.1 that all Mandarin tones might have two tonemes, and that these are the terminal nodes of the branching metrical structure normally associated with full toned syllables; when no such branching structure is present the tonemes are ultimately deleted (or were never present underlyingly in some cases). If this is right, then the four tones of Mandarin may be represented as follows:



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 as exactly that: [53] (see Chao 1968: 28) 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.

Taking the underlying representations of the tones to be as above let us continue to dissect the tonal phonology of Mandarin. The most prominent tone sandhi rule is the third-tone sandhi rule that changes the first of two third tones into a second tone. In our system this rule must effect the following change:

$$\begin{array}{c} [- \text{ Upper}] \\ \swarrow \quad \searrow \\ L \quad L \end{array} \longrightarrow \begin{array}{c} [+ \text{ Upper}] \\ \swarrow \quad \searrow \\ L \quad H \end{array} \quad / \quad \text{---} \quad \begin{array}{c} [- \text{ Upper}] \\ \swarrow \quad \searrow \\ L \quad L \end{array}$$

The change in register is clearly dissimilatory, and can be separated out as follows:

$$(1) \quad [- \text{ Upper}] \longrightarrow [+ \text{ Upper}] \quad / \quad \text{---} \quad [- \text{ Upper}]$$

Since the third tone is the only [- Upper] tone no further context is required.

The output of this rule will be [+ Upper, L L], a complex that never surfaces. Rather the second L also dissimilates by rule (2):

$$(2) \quad L \longrightarrow H \quad / \quad \begin{array}{c} [+ \text{ Upper}] \\ \swarrow \quad \searrow \\ L \quad \text{---} \end{array}$$

This will produce the right output.¹

The problems with this rule arise when we look at sequences of

third tones, and are not problems that are limited to this approach. A sequence of third tones can undergo the sandhi rule in a variety of ways, and the choice depends on the syntax, speed of utterance, and perhaps stylistic considerations (see Cheng 1973: 45-53). Quite long chains of third tones can all undergo the rule if the conditions are met, producing sequences of second tones. The problem is how to achieve this.

Take a sequence of third tones, shown in partial representation below:

[- Upper] [- Upper] [- Upper] [- Upper]

Clearly the rule as written has to apply more than once to change the first three instances to [+ Upper]. The question is how this multiple application is to be dealt with. One possibility is that it applies (simultaneously) everywhere the context is met, giving

[+ Upper] [+ Upper] [+ Upper] [- Upper]

in one step. Application of this type has been proposed by Chomsky and Halle (1968), but fails in other cases where one application of a rule feeds or bleeds an adjacent application, such as Chiluba (Howard 1972: 26). Chiluba has a rule of nasalization:

$$[+ \text{ lateral}] \longrightarrow \left[\begin{array}{l} - \text{ lateral} \\ + \text{ nasal} \end{array} \right] / [+ \text{ nasal}] \quad V_o \text{ ______}$$

When more than one lateral is present it feeds itself, as shown in:

u + dim + il + ile \longrightarrow u + dim + in + ine

There is no way of achieving this effect by a single, albeit simultaneous, application of the rule, and it seemed that some rules must truly apply to their own outputs by iterating across the word as suggested by Howard (1972) and Johnson (1970). The Chiluba rule can then be said to apply left-to-right. If simultaneous application is not sufficient, will iterative application alone suffice, and if so can some of its properties be predicted?

In order to avoid too great a weakening of the theory people tried to show that the direction of iteration was predictable from the nature of the rule, and in the majority of cases this turned out to be true. Rules normally iterate in the direction 'away from their environment', as in Chiluba, and in rules with environments on one side only this is equivalent to saying that they apply so as to maximize feeding or bleeding. Where the environment is bi-lateral these are not equivalent, and the second appears to be the right statement in such cases.

The problem arises when the rule does not feed or bleed itself, as in Mandarin. If the directionality were as predicted by the preceding remarks, the rule would apply right-to-left. This would result in the following derivation:

	[- Upper]	[- Upper]	[- Upper]	[- Upper]
1st application			[+ Upper]	
2nd application		N/A		
3rd application	[+ Upper]			

and eventually in the sequence [35 21 35 21]. There would be no way of deriving two second tones in sequence, since the rule is self-

bleeding. It is therefore necessary to say one of two things: either this is a marked rule that happens to apply left-to-right:

	[- Upper]	[- Upper]	[- Upper]	[- Upper]
1st application	[+ Upper]			
2nd application		[+ Upper]		
3rd application			[+ Upper]	

or some rules apply simultaneously (but most iteratively).

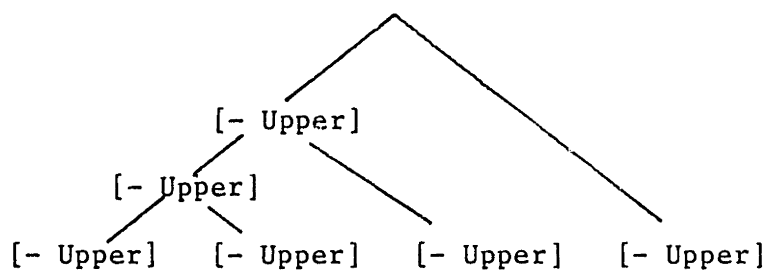
By now it should be clear that an attempt to limit multiple application to iterative rules is beset with problems, and many attempts have been made to circumvent these.

Anderson (1974) has proposed an alternative solution that allows simultaneous application subject to strict conditions on non-overlap of environments. He formulates the Mandarin sandhi rule so that it no longer either feeds or bleeds itself, and the problem then does not arise. Unfortunately Anderson's formulation of the rule leads him to postulate a further rule which has some undesirable consequences, as we will show. Anderson gives the rule as follows (p. 237):

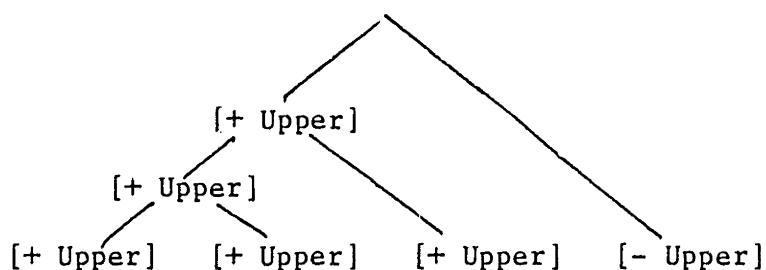
(3) L → H / . L _____ . L

where the dot indicates a syllable boundary. The right-hand environment requires only that the following tone begin with a L Tone, which therefore excludes HH, MH, and HL tones. Now the output of this rule must be indistinguishable from the underlying MH tone, so he adds a rule:

Metrical theory offers a new way of dealing with iterative rules such as vowel harmony and assimilation of other kinds. Trees are built over 'adjacent' segments of a particular kind, such as vowels. (Adjacency is defined on a projection of just those segments, so that two vowels are adjacent if they are separated by no other vowels). Nodes of this tree can then be labelled for particular features which will then percolate down to the terminal nodes. Consider how this would work for Mandarin sandhi, which dissimilates the feature [- Upper]. A tree is built over any sequence of [- Upper] in the domain; by convention the feature on a node percolates to the next higher node and so up to the top:



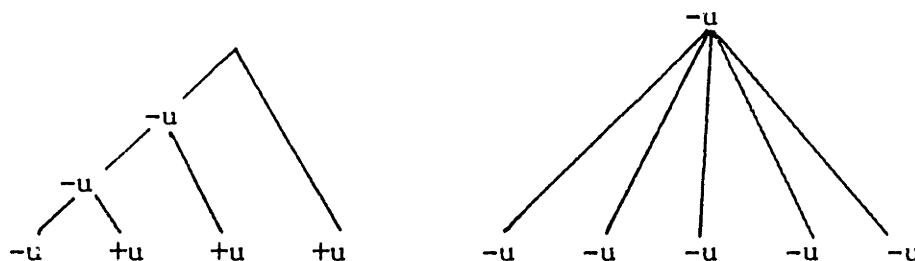
Now we need only assume that the dissimilation rule applies not to terminal nodes, but rather at the highest possible level followed again by conventional percolation (but in this case downwards):



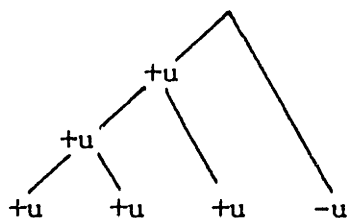
A single application of the rule will therefore effect a change in a

sequence of terminal nodes without problems. (The domain of the rule is defined by a combination of factors which we will not go into here.) The dissimilation rule will of course be followed by rule (2) which converts the sequence /LL/ to [LH] in the context [+ Upper].

This proposal deviates from previous metrical analyses (such as H&V 1978) in that the tree is binary, and yet the features percolate. The most usual assumption has been that trees are of two types: binary and polarized, or n-ary and non-polarized (i.e., percolating):



The trees given here are structurally binary, but polarized only at the highest pair of sister nodes:



The interesting question is what predictions trees of this kind make about types of assimilation and dissimilation rules. Morris Halle has argued (p.c.) that this type of tree violates the spirit of metrical theory and that straightforward simultaneous rule application of the traditional type is the right way to deal with these cases (e.g.,

jer-lowering in Russian). It is not clear to me that this extension of metrical theory is in any way more powerful than simultaneous application, and it is interesting that all major types of multiple application could be dealt with in such an extended version of metrical theory.

Let us next discuss the other tone sandhi rules of Mandarin, which are relatively straightforward. The high rising second tone becomes high level in fast speech when in non-final position and preceded by a tone that ends with [H]. This rule must follow the third tone sandhi rule, because it applies also to derived high rising tones:

	Lao ²¹	Li ²¹	mai ²¹	hao ²¹	jiu ²¹	'Old Li buys good wine'
3rd tone sandhi	Lao ³⁵	Li ³⁵	mai ³⁵	hao ³⁵	jiu ²¹	
2nd tone sandhi	Lao ³⁵	Li ⁵⁵	mai ⁵⁵	hao ⁵⁵	jiu ²¹	

The rule is clearly one of assimilation, and applies only when the tones are dominated by a W node of the stress tree -- that is, non-finally:

(5)



Next we must deal with the pre-pausal form of the third tone. We have already noted (section 1.2.1, 1.5) that in order to derive the phonetic rise we must assume that a H tone is inserted after the third

tone in two environments: absolute final position and before a neutral tone. The question is how to unify these. We suggested two ways in section 1.2.2: either the environment is "when no tone directly follows in the domain" or it makes reference to stress, since both cases are dominated by S. It turns out that the second is the correct statement, for the following reasons. Metrical trees are built at only two levels in Mandarin: the foot and the phrase. So-called word stress is nothing but phrase stress on an isolated word. This can be seen from the fact that within a phrase there are no stress contrasts caused by different word structure. Consider the two verb phrases:

[mai²¹ hao²¹] jiu²¹ 'have bought wine'
 mai²¹ [hao²¹ jiu²¹] 'bought good wine'

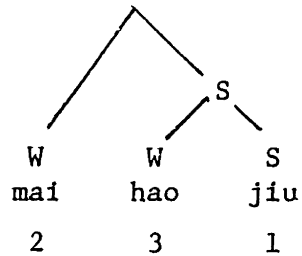
When these are spoken at fast speed both the first two syllables undergo third tone sandhi, so they are tonally indistinguishable:

mai³⁵ hao³⁵ jiu²¹⁴

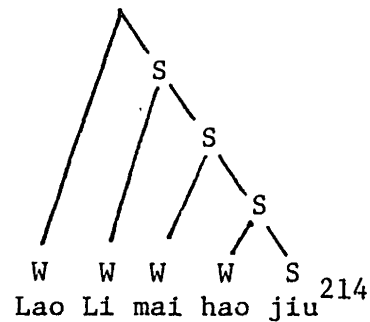
Having guaranteed a single phonological phrasing, are there still distinctions of stress? If word trees are built first we should get a difference in subsidiary stresses as shown:



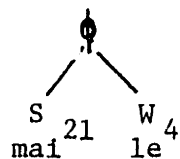
In fact they are both 231, the pattern given by a single level of phrase stress:



Returning then to the context for the rule inserting H tone after the [- Upper, LL] third tone, the first environment, absolute final position, is dominated by S in the phrase tree:



The second environment, before a neutral tone, is dominated by S in the foot tree:



So if we state the environment as 'when immediately dominated by S', it will produce the right result.

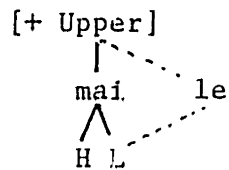
The last purely tonal phenomenon is neutral tone. We have argued earlier that there is a rule that deletes the Tone in some environment, and suggested that the relevant environment is when dominated by a non-branching structure at the foot level (i.e., in W position in a foot):

$$(6) \quad T \longrightarrow \phi \quad / \quad \begin{array}{c} \phi \\ | \\ W \\ | \\ \text{---} \end{array}$$

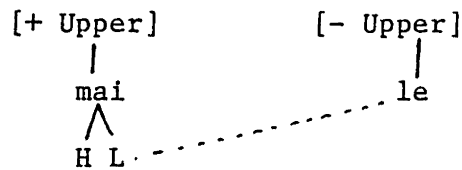
It is also true that Register is always [- Upper] in this environment, so we need a rule:

$$(7) \quad R \longrightarrow [- \text{Upper}] \quad / \quad \begin{array}{c} \text{---} \\ | \\ W \\ | \\ \text{---} \end{array} \left. \vphantom{\begin{array}{c} \text{---} \\ | \\ W \\ | \\ \text{---} \end{array}} \right] \phi^2$$

This second rule will change the underlying Register of some morphemes to [- Upper], but it will also act on morphemes with no underlying Register or Tone of their own. For example, consider the suffix le (perfective aspect) when added to a high falling toned morpheme:



The WFC condition will automatically cause spreading of both Register and Tone, but rule (7) will then act to produce:



Note that rules (6) and (7) must follow third tone sandhi rules (1) and (2), since the tones that trigger sandhi may subsequently be deleted.

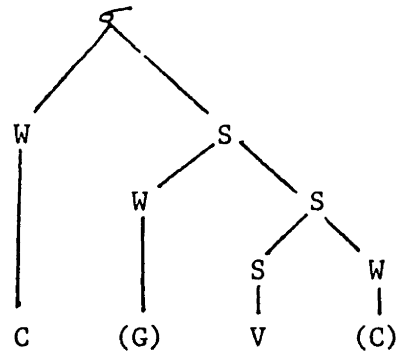
The remaining tonally related phenomena involve segmental changes under tonal conditioning. The first has already been mentioned in section 2.1.2: before a neutral tone unaspirated stops become voiced, and syllable final consonants resyllabify under certain circumstances. Let us now spell out the rules needed for Mandarin, which differ from those for Amoy.

Resyllabification takes place when the first syllable ends in a consonant and the second starts in a vowel provided the second syllable is in neutral tone. Such cases are limited to sentence particles, as far as I know, although it should be possible to find others. For example, Chao (1968: 803) gives:

人 呀	rén a	→	rén (n)a
娘 呀	niáng a	→	niáng (ng)a

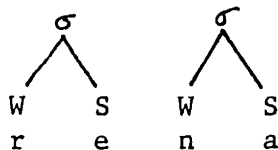
This can easily be understood if we assume that the syllabic template for Mandarin requires an initial consonant wherever possible, but that the final consonant is optional:

(8)



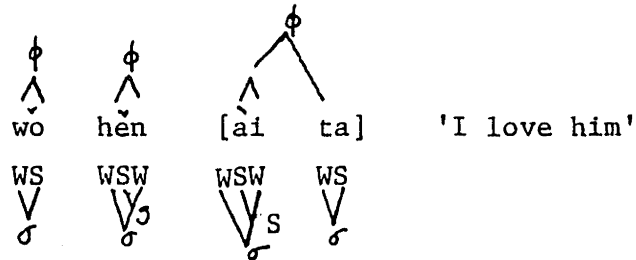
Kiparsky (1979) has argued that this is a universal fact that is explained by the fact that the minimal syllable structure has only the single branching $W \quad S$, which will produce a CV syllable when mapped onto the sonority hierarchy. Note by the way that the structure given above for Mandarin departs from Kiparsky's possible syllable types in that the glide is put in the rhyme, as is traditional in Chinese phonology, rather than in the initial. Nothing crucial depends on this decision here, but the reader should be aware that the glide certainly behaves as part of the rhyme in word games, just as the [y] does in English Pig Latin where cute goes to [yũwtkēy] (McCarthy 1979: 19). (Note that the underlying medial glide in English, [w], patterns as part of the initial: quick goes to [Íkwēy]).

Given the template above, the syllabification of a string like rén a will be as follows:



Now this only happens when the second syllable is in neutral tone, so apparently the domain of syllabification is the stress foot (see

section 2.1.2), which explains why no such resyllabification takes place between hèn and ài in a string like:



We will now turn to the voicing of unaspirated initial consonants in such syllables. Cheng (1973: 82) points out that aspirated initials often lose their aspiration in the same environment, and suggests that rather than a voicing rule what is needed is a laxing rule in neutral toned syllables, which we would write:

$$(9) \quad C \longrightarrow [- \text{tense}] / \underline{\quad \begin{array}{c} W \\ | \end{array} \quad}$$

However, Cheng also points out that there is a general process of vowel laxing in neutral toned syllables (given in phonetic transcription, not pinyin):

$$\begin{array}{l} k\bar{\epsilon} \quad k\bar{\epsilon} \longrightarrow k\bar{\epsilon} \quad k\bar{\epsilon} \quad \text{'elder brother'} \\ x\acute{u}i \quad l\acute{a}i \longrightarrow xu\acute{e}i \quad lai \quad \text{'to come back'} \end{array}$$

and suggests that in fact the entire neutral toned syllable can be seen as becoming lax:

$$(10) \quad \sigma \longrightarrow [- \text{tense}] / \underline{\quad \begin{array}{c} W \\ | \end{array} \quad}$$

The feature [- tense] then percolates down to all segments. This laxing rule can feed other rules. In particular, low lax vowels reduce to schwa [ə], and can subsequently assimilate in roundness to an adjacent glide. See Cheng (1973: 73-83) for a full analysis.

There are two other vowel changes that can be conditioned by specific tones. In neutral toned syllables after fourth tone syllables vowels can devoice. The context is very easily captured in our system since after a fourth tone the neutral tone is [- Upper, L] whereas after all the other tones it is [- Upper, H]:

$$(11) \quad V \longrightarrow [- \text{voice}] / \begin{array}{c} \text{W} \\ | \\ [- \text{Upper}] \\ | \\ \hline | \\ \text{L} \end{array}$$

The last process involves vowel alternations between the mid and high vowels in the following finals:

<u>Third, Fourth and Neutral</u>	<u>First and Second</u>
iou	iu
uei	ui

Unfortunately it is not immediately apparent what the context for this change is. Assuming that the rule is one of raising, we may write:

$$(12) \quad V \longrightarrow [+ \text{high}] / \begin{array}{c} \hline \wedge \\ \text{T} \quad \text{H} \end{array}$$

This will have the desired effect, but it is not the kind of rule one

expects to find for the following reason. Part of the environment is a H tone, but it must be the last tone associated with that syllable, since the HL fourth tone does not trigger the raising rule. We have argued in section 4.4 that the domain of a tone in such instances should be the later part of the syllable and not, as in this case, the nucleus. The type of rule we would expect to find would be:

$$(13) \quad V \longrightarrow [+ \text{high}] / \begin{array}{c} [+ \text{Upper}] \\ | \\ \hline \end{array}$$

since the domain of Register is the whole syllable. But if this is the rule, then we must reconsider the underlying representation of the fourth tone, since it would have to be [- Upper, HL] instead of [+ Upper]. This would have a number of undesirable consequences. First it would fail to explain the [53] form found on the first of two fourth tones; second, fourth tones would now be subject to the third tone sandhi rule which dissimilates Registers, so that third tones would, wrongly, change to second tones before fourth tones as well as other third tones. It would therefore be necessary to completely revise the analysis of sandhi. I would like to suggest another possibility. The alternations in question are rather marginal, and many speakers do not have them at all. It may be that they are disappearing from the language precisely because the rule is so highly marked; certainly it is usually the older sources that mention the alternations. For this reason I conclude that the balance of the evidence favours representing the fourth tone as [+ Upper] Register, and accepting the deviant rule (12). This may not be as bad as it

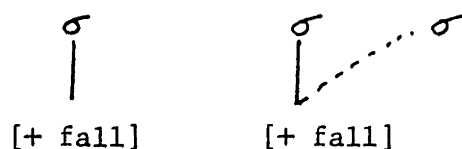
seems: note that the preceding glide is also part of the rhyme, so the affected V is towards the end of the rhyme, as is the conditioning H tone.

This concludes the discussion of the tonal phonology of Mandarin. As far as I am aware all tonally related phenomena have been dealt with.

5.2 Shanghai

The data on which the following analysis is based are taken from Zee and Maddieson (henceforth Z&M) (1979) except where otherwise stated. Their article is the only source of phonetic data on Shanghai as far as I know, and provides an excellent data base to work from. Some aspects of the phonological analysis that follows are also due to Z&M, but in many respects I depart from their analysis, and of course the Register system used here is not used in their paper.

The general conclusion reached by Z&M is that Sherard (1972) was largely right in claiming that the tone contour of polysyllabic words depends only on the tone of the initial syllable, but wrong in implying that the contour (as a unit) of the first syllable was spread over the longer word. That is, a syllable which is falling toned in isolation produces a bisyllabic word with high tone on the first syllable and low tone on the second. There is no fall on the first syllable, a fact which would be inexplicable if a feature [+ fall] were attached to that syllable and spread over the rest of the word:



Instead the sequence /HL/ is realized on one syllable as a fall, but on two syllables as a sequence of high tone followed by low tone. Thus far Z&M are absolutely correct, but their analysis of the detailed facts is unfortunately distorted by too great a focus on minor phonetic details so that certain broad trends escape their notice. I will first present my analysis, and then point out where it differs from Z&M.

On the following pages I reproduce the most important data from Z&M, showing F_0 tracings of monosyllables and polysyllabic words of various lengths.

Two assumptions underlie what follows. The first is that it is desirable to reduce the tonal inventory as far as possible, and the second is that there is a link between tonal Register and voice quality, so that those tones which Sherard describes as occurring with murmured release, tones C and E, belong to one Register, and the others to the opposite Register; specifically, I take it that murmur, which is known to be associated with low pitch, is [- Upper] Register, while the other tones are [+ Upper].

With these assumptions in mind, let us examine the five tones of Shanghai (see figures 1 and 2). Two of them, D and E, end in a glottal stop, and the sharp fall in pitch observed at the end of these syllables irrespective of what tones they bear can be attributed to this and will therefore be disregarded in what follows.

Tone A is, by hypothesis, [+ Upper] Register, and it is clearly falling (all authorities agree on this). We will therefore assume it to be [+ Upper, HL].

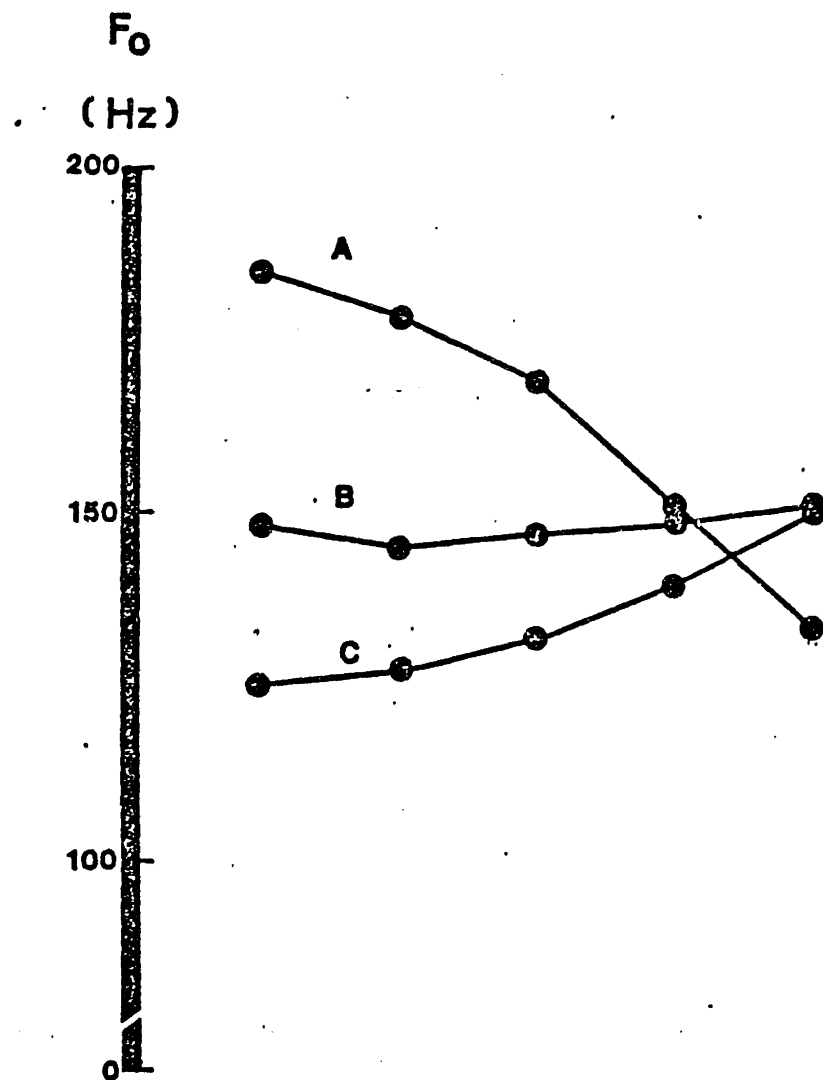


FIGURE 1. Pitch contours of normalized duration of syllable types A, B, and C.

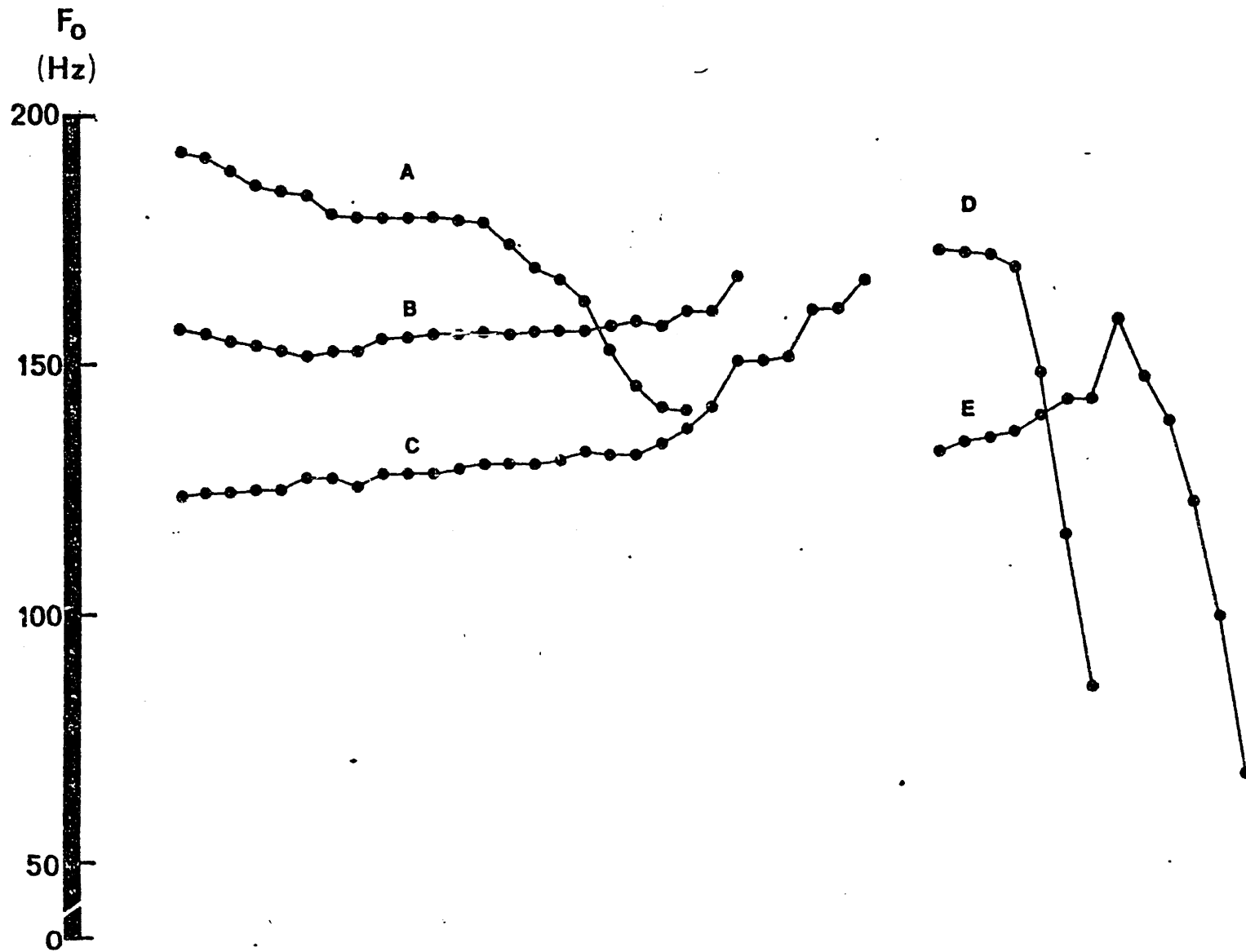


FIGURE 2. Pitch contours on syllable types A, B, C, D and E.

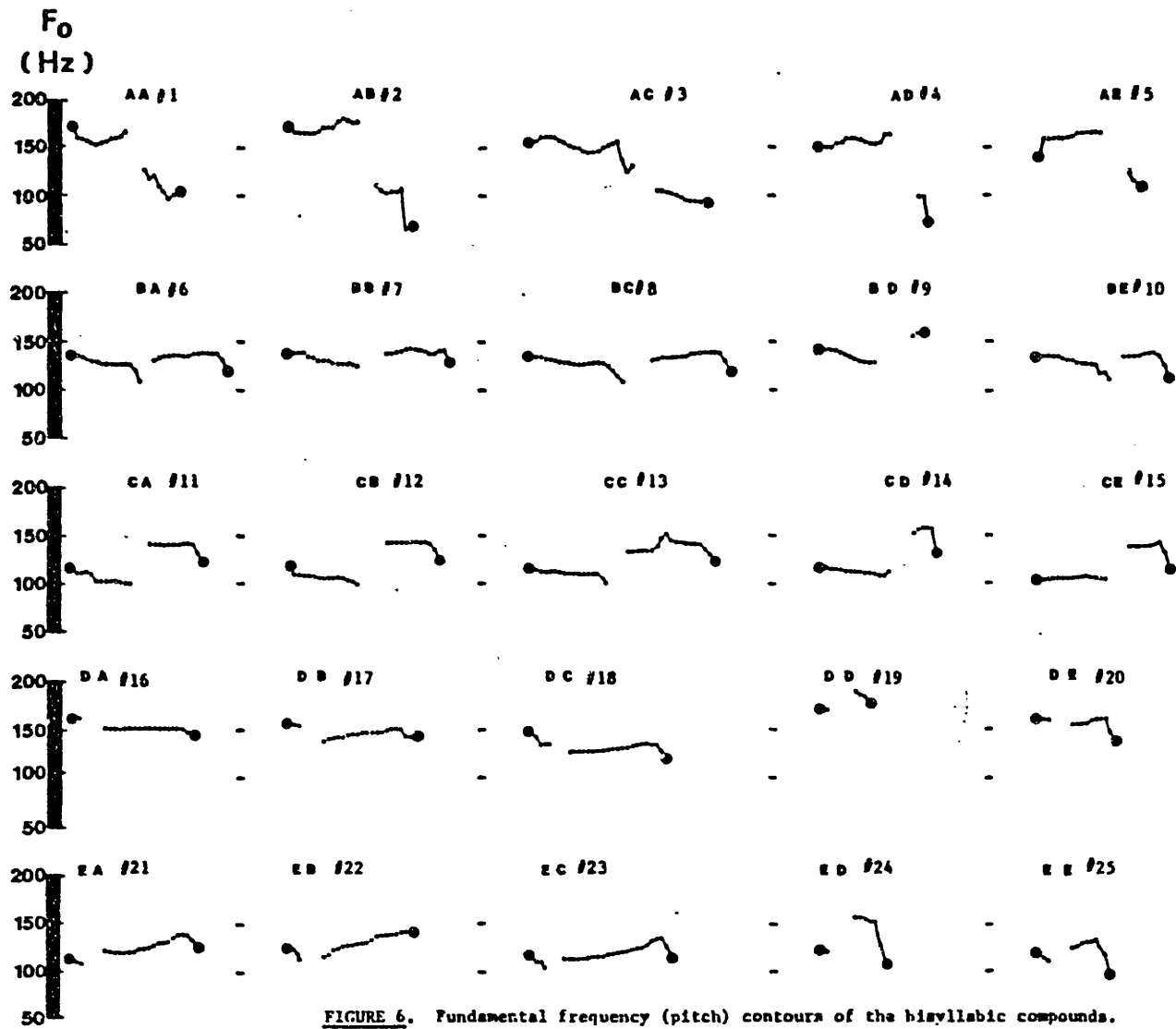


FIGURE 6. Fundamental frequency (pitch) contours of the bisyllabic compounds.

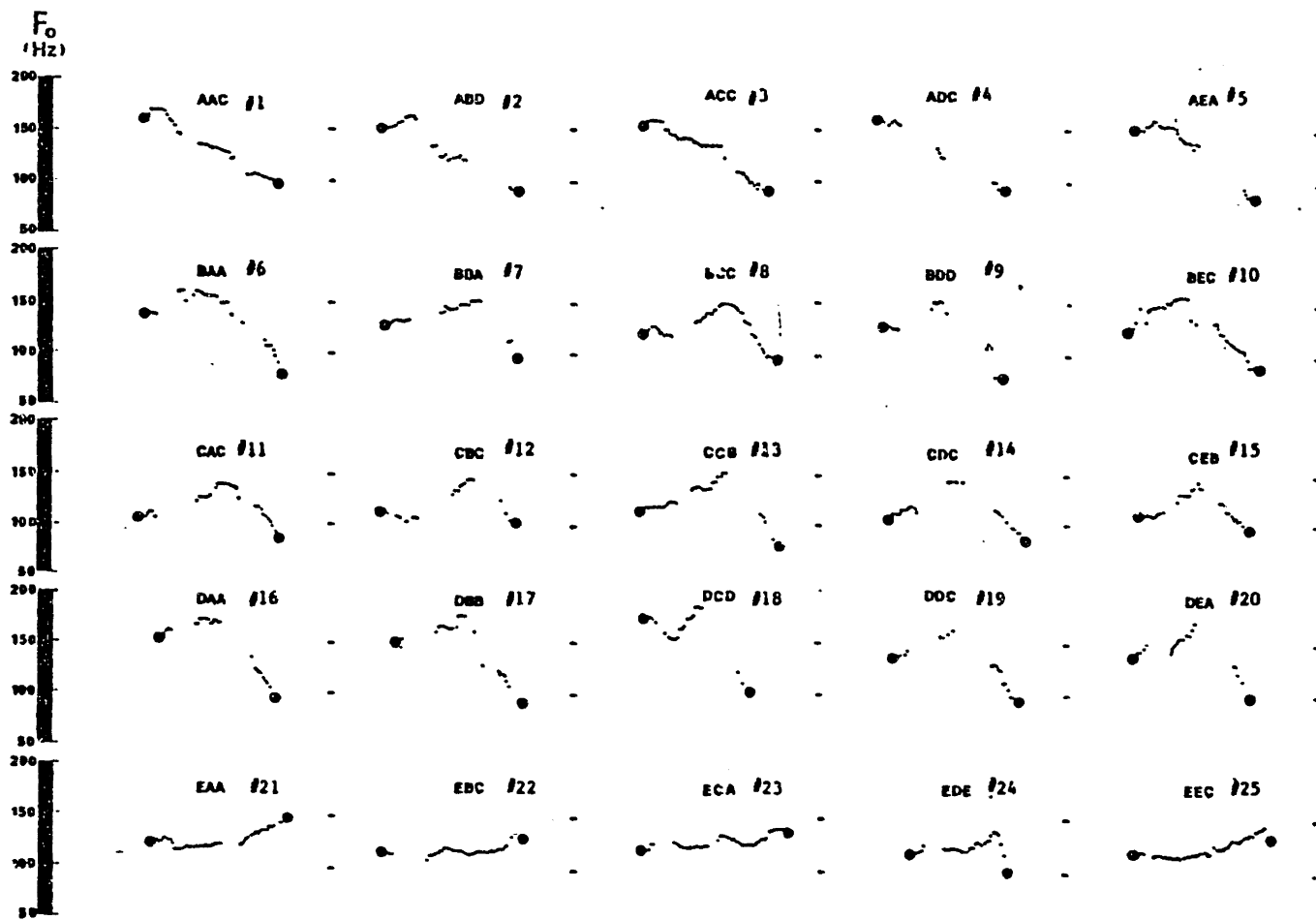


FIGURE 7. Fundamental frequency (pitch) contours of the trisyllabic compounds.

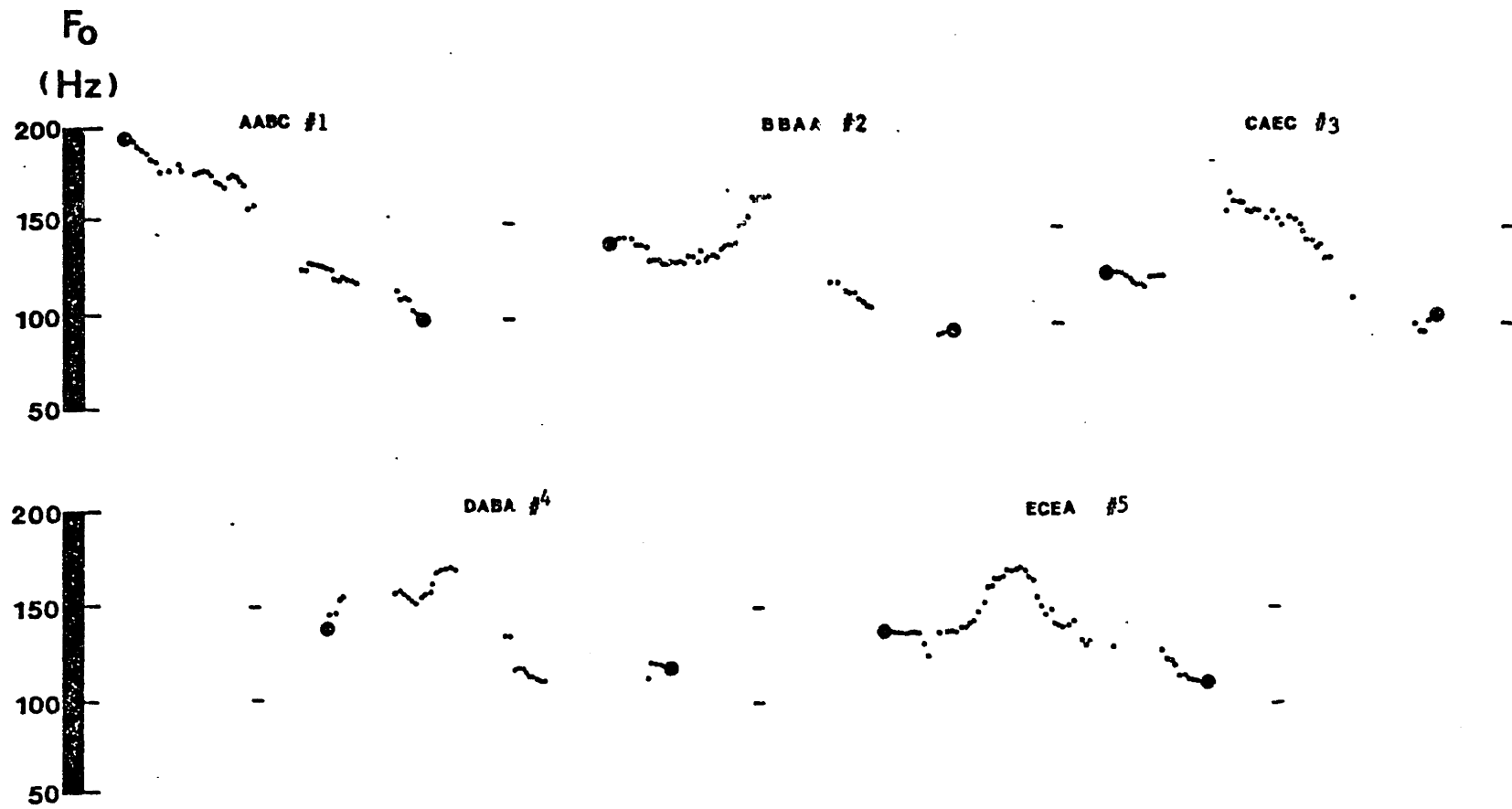


FIGURE 8. Fundamental frequency (pitch) contours of the quadrisyllabic compounds.

Tone B is also [+ Upper], and basically level. It is nothing like as high as the starting point of tone A, and is therefore presumably [+ Upper, L]. Z&M consider it rising, but the rise is no more than 3 or 4 Hz, compared to the 25 Hz of a true rising tone like tone C.

Tone C is clearly rising, and since it is associated with murmur it must be [- Upper, LH].

Tone D is level, and very much lower than the starting point of tone A. For this reason I take it to be [+ Upper, L], although it is certainly somewhat higher than tone B, with which I am now claiming it is tonally identical and differentiated only by the final glottal stop. We will see, however, that as the first element in a polysyllable (when the glottal stop is lost) it produces remarkably similar contours to tone B which supplies a second reason for taking it to be identical.

Lastly, tone E has a slight rise, and since it, like tone C, is associated with murmur I take it to be the variant of C in the presence of a glottal stop, [- Upper, LH].

We thus have a tonal inventory of three tones (as opposed to Z&M's 4):

A	[+ Upper, HL]
B, D	[+ Upper, L]
C, E	[- Upper, LH]

In order to account for the monosyllabic contours only one additional note is necessary. Tone A, which I have claimed to be [+ Upper, HL], actually falls right down to the lowest possible pitch. This is

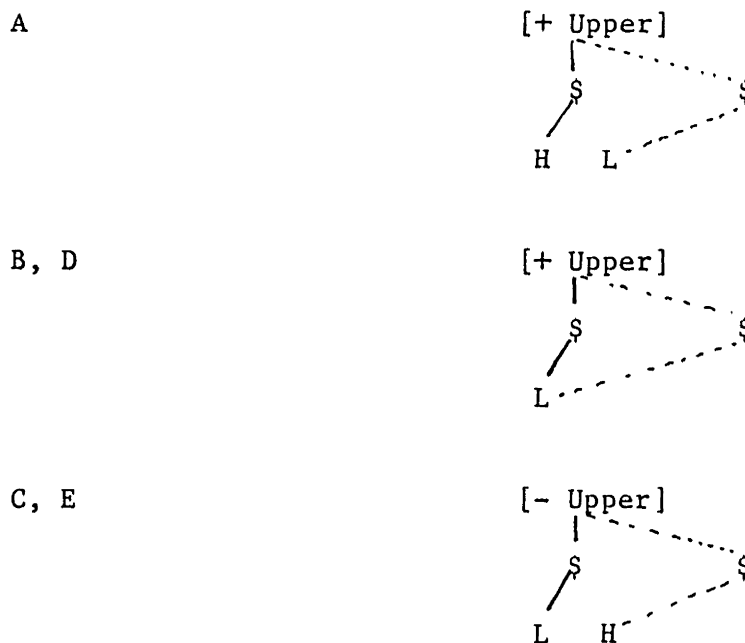
remarkably reminiscent of the Mandarin fourth tone, which is also [+ Upper, HL], but also falls lower than expected in isolation. We will see that just like the Mandarin tone, Shanghai tone A falls only to the expected mid-level when something else follows. We therefore take it that the extra fall word-finally is, like in Mandarin, nothing but a non-phonological prolongation of the essential tone.

The more interesting cases involve longer words. First consider the bisyllables. We have already said that the tone of all but the first syllable is irrelevant (there is one exception to this, when the second syllable is tone D, and we will return to this in a moment).

Suppose that Shanghai has the following language-specific rule of association:

- (1) Associate the first (leftmost) tone with the first (leftmost) syllable.

This is a very common type of initial association rule (see Clements and Ford 1979: 181) and indeed in some languages (such as Slavic) lexical entries may already include this single association line. After initial association has taken place the WFC will come into operation for the remaining unassociated syllables, and the results are shown below (dotted lines are those inserted by the WFC):

Bisyllabic WordsFirst Syllable

If we now compare this with the phonetic facts in Figure 6, the contours are exactly as expected with the following qualifications. There is a slight, abrupt fall at the end of nearly all words, which I take to be automatic and uninteresting. Tone A is lower at the end than expected, just as it is in monosyllables, and presumably as a result of the same process. This is to be expected, since it is still the last tone in the word. Tones B, D and C are exactly as expected, but tone E is apparently subject to a further rule that spreads the L tone rightwards resulting in a low rising rather than a mid level tone on the second syllable. We will postpone formulation of this rule until we have examined the three and four-syllabled words, which differ in certain respects.

The exception to all the above is when the second syllable is

in tone D. For reasons that I do not understand it appears that the tone D is not deleted in bisyllabic words, since the second syllable always has a mid-high level pitch like the isolation contour of tone D. Assuming that the tone remains behind, we have the following forms:

Bisyllables Ending in Tone D

First Syllable

A	[+ Upper] \$ / \ H L	[+ Upper] \$ L
B, D	[+ Upper] \$ L	[+ Upper] \$ L
C, E	[- Upper] \$ L H	[+ Upper] \$ L

Notice that certain tones remain unassociated. Clements and Ford (1979: 186) have proposed a more precise but somewhat weaker form of the WFC that allows exactly this: unassociated tones. They further suggest that such tones will simply fail to be realized phonetically unless subsequently associated by language specific rules. We are suggesting that this is exactly what happens here, since if the spare tone were associated with either syllable we would expect to find a contour tone on one or other syllable, whereas in fact both

are quite level. Once we accept this weakened WFC, the forms of bisyllables ending in tone D are accounted for without any trouble. Notice again that tone A falls further than would be expected, for the same reasons as before (the output before tone D or before any other tone being identical both phonologically and phonetically).

Once a word is longer than two syllables, something rather different happens. No matter what the underlying tone of the first syllable is, the third and subsequent syllables are realized on a low level. There is (as usual) one exception to this, trisyllabic words beginning in tone E, and we will return to this later. The simplest way to capture this, as Z&M suggest, is to assume that a low tone is inserted on the third syllable. This will produce the following trisyllabic forms after initial association and operation of the WFC:

Trisyllabic Words

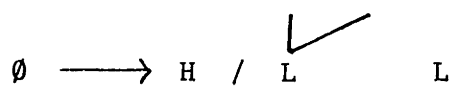
First Syllable

A	[+ Upper] \$ / \ H L	[- Upper] \$ L
B, D	[+ Upper] \$ L	[- Upper] \$ L
C	[- Upper] \$ / \ L H	[- Upper] \$ L

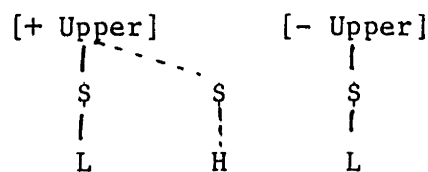
The preference for spreading from the left rather than from the right is also in accordance with Clements and Ford's formulation of the WFC (Convention 3, p. 186).³

Comparing these with the phonetic facts in Figure 7 we notice that tone A no longer falls extra-low. This is exactly as predicted, since it is no longer the final tone in the word. Tone C is precisely as expected, but one further rule is needed for tones B and D. Notice that there is a clear rise on the second syllable in both cases, and this rise is not predicted by the forms above. We therefore need a rule which inserts a H in a sequence of L's:

(2) H Insertion



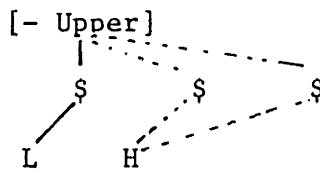
In the output the tones must then be associated one-to-one as follows:



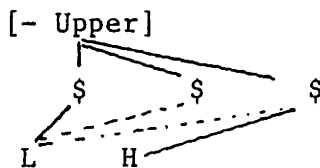
It must therefore be the case either that H-insertion precedes operation of the WFC, or that by convention it erases the preceding association line between the first L and the second syllable. I will take it that the second is the case: in other words, substitution is preferred to multiple attachment wherever possible.

I mentioned that tone E is exceptional. The exceptionality lies

in the apparent absence of the inserted low tone in trisyllables beginning with tone E. As with the aberrant behaviour of tone D in bisyllables, this is a complete mystery, but I will assume it is the case. (One could of course write a deletion rule, but it would be descriptive only and therefore quite uninteresting.) The words beginning in tone E will then have the following representation:

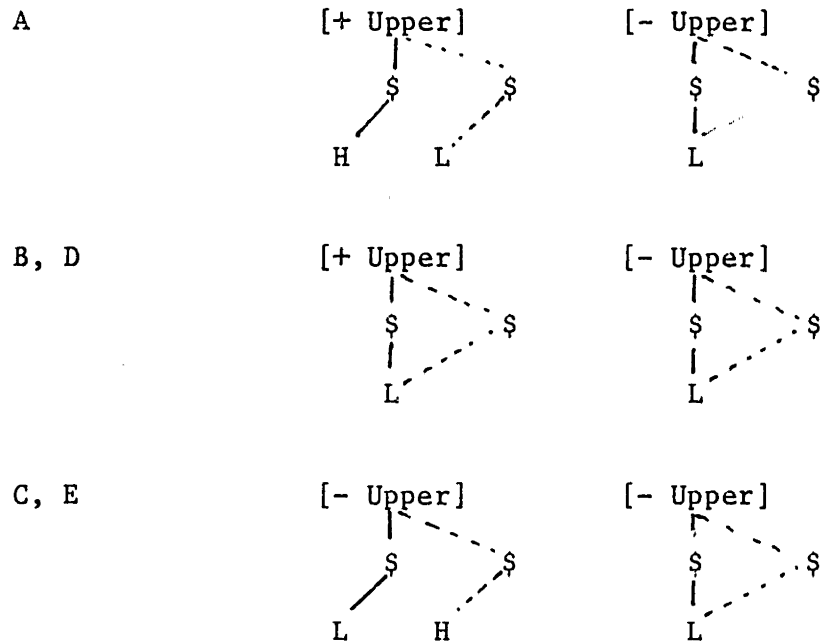


This is still not right, however. In bisyllables the L of tone E spreads rightwards across the word, and the same happens here, producing the following:

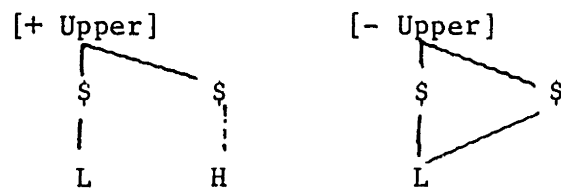


After discussing the four-syllable forms we will finally be ready to state the rule that does this.

Let us lastly deal with the four-syllabled forms shown in Figure 8. They turn out to be quite straightforward, since even tone E has the inserted low tone on the third syllable:

Four-Syllabled WordsFirst Syllable

As in the three-syllabled words, tone A surfaces exactly as expected, with no specially low fall on the non-final L tone. Words beginning in B and D undergo H-insertion as expected, to produce:

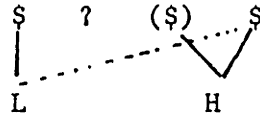


Words beginning in C and E both surface exactly as expected, with a rise on the first two syllables (it is somewhat higher than might be expected in the case of tone E, but I doubt that this is significant, and only a single token is given so it may not even be typical).

It remains only to formulate the rule that spreads low tone rightwards from tone E in bisyllabic and trisyllabic forms. The rule

must spread the L tone rightwards up to the end of the word, but it only applies if it is the last L tone in the word (i.e., not in quadrisyllabic forms):

(3) L Spreading

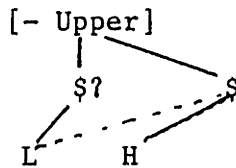


(where the rule inserts the dotted line(s).)

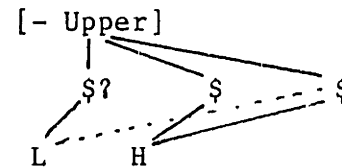
The glottal stop restricts the rule from applying to tone C, and is later deleted.

The rule will apply as follows:

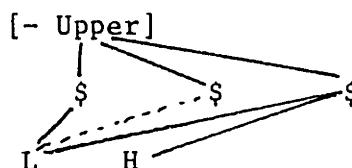
Bisyllables



Trisyllables



One of the requirements of the WFC is that association lines cannot cross. If the output of a rule violates this, by convention it is the old association lines that are erased, rather than the new one(s) inserted by the rule. The final output of the trisyllabic form will therefore be (after deletion and reapplication of the WFC):



To summarize, Shanghai has an inventory of three tones, two of which can occur on stopped or unstopped syllables while the remaining one occurs on unstopped syllables only. One of the three tones is [- Upper] Register and produces murmur on the word-initial consonant. With one exception, tone D in word-final position, the Register and Tone of non-initial syllables are deleted. A low tone, [- Upper, L], is inserted on the third syllable of long words. Shanghai has the initial association rule "Associate the first tone with the first syllable," and this, followed by the WFC as reformulated by Clements and Ford, will account for all the observed facts with the following additions:

- (i) Tone A falls extra low in word-final position; this is a surface fact with no phonological repercussions.
- (ii) L-spreading (rule (3))
- (iii) H-insertion (rule (2))
- (iv) No low tone is inserted after tone E in trisyllables, for some reason.

A summary chart appears overleaf.

Let us now compare this with Z&M's solution. They have an inventory of four tones because they take tone B to be rising. They also delete non-initial tones with the exception of tone D, but since they do not consider the lowness of the end-point of tone A extra-systemic they are forced to delete tone D after tone A but not elsewhere, a slight additional complication (Z&M rule (8)). Furthermore, because they assume that association precedes deletion,

SHANGHAI

	AX	AD	BX DX	BD DD	CX EX	CD ED	AXX(X)	B D XX(X)	CXX(X) EXXX	EXX
H insertion (2)										
L Spreading (3)					<u>EX only</u> 					
H	—	—			CX		—	—		
M	—	—	—	—			—	—	—	—
L	—	—					—	—	—	—
	But actually falls to L word-finally				EX					

rather than the reverse, they require a de-association process to convert



As a result of taking tone B to be rising, they are forced to a rule of H-lowering that reduces it to an almost-level tone in certain contexts. What is more, they resort to a global conditioning for this rule (Z&M rule (9)'). As a further consequence of associating all tones before deletion they never arrive at an unassociated (and therefore unrealized) tone in bisyllables before tone D; as a result a special rule of Tone Deletion is needed for these cases (Z&M rule (11)). Their analysis, like mine, recognizes a rule of low tone insertion on the third syllable, and the exceptional behaviour of trisyllables beginning in tone E (Z&M rule (12)), and they include a close equivalent of my rule (3) of L-spreading, although they treat it as an initial association rule. The rise in pitch that we have analyzed as resulting from the dissimilatory rule of H-insertion, rule (2), has two different sources for Z&M. The rise on tone B is simply a manifestation of its underlying form (which is flattened elsewhere by H-lowering) and the rise on tone D is achieved by lowering the first of these two -- a process they take to be phonetic, and give no rule for (see p. 126). It should be clear, I think, that whereas my solution involves only two language specific rules (apart from the initial association rule that all tone languages must have) and two exceptions, Z&M need the same exceptions plus a plethora of rules,

special associations and de-associations. Furthermore, their tonal inventory has four members as opposed to my three.

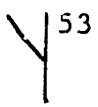
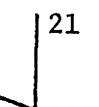
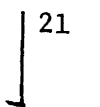
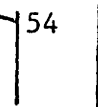
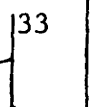
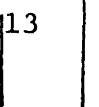
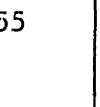
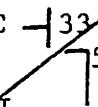
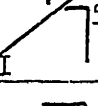
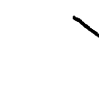
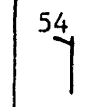
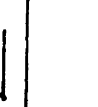
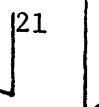
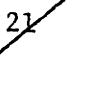
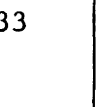
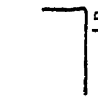
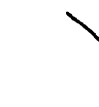
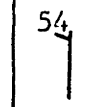
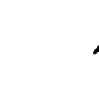

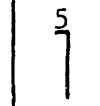

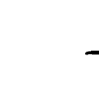
I conclude that this analysis of Shanghai strongly supports the Register proposal, and also provides evidence in favour of several aspects of Clements and Ford's re-statement of the WFC.

5.3 Amoy

The most interesting feature of the Southern Min languages like Amoy is that every tone has two forms, one of which occurs pre-pausally and before neutral tone -- the so-called citation form -- and the other of which occurs elsewhere -- the sandhi form. There may also be other variants that occur in special environments, and in the case of Amoy itself there are two of these: one appears in the first syllable of triply reduplicated adjectives, and the second appears before a particular suffix. The four forms of each tone are summarized on the chart overleaf. In two cases the sandhi form (E2) is different in two sub-dialects, Coastal and Inland (Hsieh 1976), and both forms are given. The remaining data are taken from R. L. Cheng (1973), and refer to the Inland dialect.

The first question one might ask is which of the two forms -- citation or sandhi -- is underlying. The traditional Chinese approach has always been to assume that the citation form is underlying, but there is no a priori reason why this should be the case. Furthermore it is easy to imagine a situation in which reanalysis had taken place so that some dialects had the citation form as underlying while others had the sandhi form as underlying; it is therefore not necessary to

TONES OF AMOY CHINESE (adapted from R. Cheng, 1973)

	2	3	4	8	7	5	1
E1	 53	 21	 21	 54	 33	 13	 55
E2	 33  55	 53	 54	 21	 21	 21	 33
E3	 55	 53	 54	 35			
E4		 55	 5	 3	 33		

Environments

E1: pre-pausal, or before neutral tone. Citation form.

E2: non-pre-pausal (except E3, E4). Sandhi form.

E3: first syllable of triply reduplicated adjectives.

E4: before the suffix a⁵³ (diminutive, co-ordinative, or nominalizing).

Notes: (i) 4 and 8 are stopped tones (velar or glottal). If the stops are lost they merge with 3 and 7 respectively.

(ii) Tone 3 in E1, and tone 7 in E2 are shown with a smaller fall than Cheng gives, in line with Hsieh's data, and HYFYGY.

(iii) Tones 2 and 5 are given with both Inland (I) and Coastal (C) variants in E2.

insist that the same form be underlying for all dialects. No argument can be made from language acquisition, since (except for nouns) it is not clear that the primary data will usually be in citation form. That said, the decision as to which is basic cannot be made on universal grounds but must be based on language-specific data. In Amoy it is clearly simpler to assume that the citation form is basic because there is neutralization in sandhi form.

In the Inland dialect tones 5 and 1 are neutralized to [33], while in the Coastal dialect tones 5 and 7 (including tone 8 that have lost their stop) are neutralized to [21], and tones 2 and 1 are neutralized to [33]. We will therefore assume that the forms in E1 are basic, since the alternative would be to postulate more abstract underlying forms, and there is no justification for that here.

The second question that must be answered is what constitutes the context for the changes from citation to sandhi forms. The environment is usually stated as the complement of 'pre-pausally, or before a neutral tone'. Notice that we have encountered this environment before when discussing the Mandarin rule that inserts a H tone after the low third tone (see 5.1). We concluded that the right way to capture this apparent disjunction was to condition the rule(s) to apply to tones dominated by a S node. Since Amoy has essentially the same stress facts as Mandarin (see 2.1.2) the same generalization can be used here in reverse: the sandhi rules will apply 'when dominated by W', and since this environment is shared by all the rules that follow, it will not be stated separately in each rule.

We are now in a position to begin an investigation of exactly what rules effect the change from citation to sandhi forms. Evidence of three kinds will bear on this investigation. The first kind of evidence (and the only kind used by many past investigators) is the surface forms of the tones. If a tone is described as 55 then it is clearly [+ Upper, HH], but if it changes to 33 it is not possible from the form alone to tell whether it has changed Register, to [- Upper, HH], or Tone, to [+ Upper, LL]. The surface forms will therefore provide an unambiguous clue to the phonological form in only a few cases. The second type of evidence comes from the subsequent behaviour of the sandhi forms when in more specific environments such as E3 and E4. In sections 1.4.3 and 4.2.1 we argued that in environment E4, before the suffix a⁵³, the Tones of a morpheme are deleted but its Register remains behind. The level of the output tone is thus an amalgam of the underlying Register of the sandhi tone, and the H Tone that spreads leftwards from the suffix in accordance with the WFC. If this analysis is right it gives us a way of deciding on the Register of the sandhi tones: those that surface as [33], [- Upper H], before the suffix a⁵³ are [- Upper], while those that surface as [55], [+ Upper, H] are [+ Upper]. This immediately resolves our earlier dilemma as to whether tone 1, which is [55], [+ Upper HH], changes Register to [- Upper, HH], or Tone, to [+ Upper, LL]. The former is the right analysis, since the [33] output behaves like a [- Upper] Register tone before a⁵³.

The third type of information is the difference between the Inland and the Coastal dialects. The differences are as follows:

E1	E2	
	Inland	Coastal
13	33	21
53	55	33

If we assume that the tone given as 21 is phonologically level, and the fall is a phonetic detail, then both dialects have two level tones but those of the Inland dialect are higher than those of the Coastal dialect. There are two ways in which we might account for this dialect difference. The first hypothesis would be that both dialects have one H and one L tone, and the difference lies in the fact that the Inland dialect has both Upper register, while the Coastal dialect has both Lower register. However this cannot be right, because we have already seen (see chart on p. 320) that the behaviour of the tones in E4 shows them to be of different registers. The second hypothesis would be that both dialects have one Upper and one Lower register tone, and the difference lies in the fact that the Inland dialect has both H, whereas the Coastal dialect has both L. This hypothesis is borne out by the subsequent behaviour of the tones in E4. (At this point it should be noted that the data available to me on the Coastal dialect makes no mention of the forms found in E3 and E4, giving only E1 and E2. The analysis given here predicts that the two dialects will not differ in E3 and E4: should that be false, some changes would have to be made. It could be, for example, that the two tones (tones 2 and 1) which

have the contour 33 in E2 in the Coastal dialect turn out to merge in E3 and E4 also -- contrary to the predictions of this analysis.)

The alert reader will notice that all of the tones have underlying forms consisting of two Tones, and we suggest that, just as in Mandarin, this provides the branching structure on which the stress trees are built. Since some output tones have only one Tone the stress trees must be built on underlying forms; in any case, if they are to be used to condition rule application they must of course be constructed beforehand.

Now let us consider the changes from E1 to E2. First consider tone 3. Since it changes from low level to High falling, it has clearly changed register. However, not every tone changes register. So consider the contrast between tone 3 and tone 2 (Inland dialect):

	E1		E2
tone 3:	21	[- Upper]	53 [+ Upper]
		$\begin{array}{c} \diagup \quad \diagdown \\ L \quad L \end{array}$	$\begin{array}{c} \diagup \quad \diagdown \\ H \quad L \end{array}$
tone 2:	53	[+ Upper]	55 [+ Upper]
		$\begin{array}{c} \diagup \quad \diagdown \\ H \quad L \end{array}$	$\begin{array}{c} \\ H \end{array}$

The difference between these tones is contour versus level, so let us hypothesize that there is a rule of register change of the following form:

(1) Register Switch

$$[\alpha \text{ Upper}] \longrightarrow [-\alpha \text{ Upper}] / \begin{array}{c} \diagup \quad \diagdown \\ \beta T \quad \beta T \end{array}$$

This rule alone accounts for a good part of the observed alternations.

So consider tones 1 and 7:

	E1		E2
tone 1:	55	[+ Upper] \wedge H H	[- Upper] 33 \wedge H H
		<u>register switch</u> →	
tone 7:	33	[+ Upper] \wedge L L	[- Upper] 21 \wedge L L
		→	

Now let us turn to the two dialects. The tones that have different reflexes are contour tones, [13] and [53], so register switch will not apply. The dialect difference can now be seen to be a difference in 'dominance' that results in survival of H's in Inland Taiwanese, but survival of L's in Coastal Taiwanese.

Let us formulate the rule as shown

(2) Recessive Deletion

$$-\alpha H \longrightarrow \emptyset \quad // \quad \alpha H$$

Condition: Inland dialect, $\alpha = +$

Coastal dialect, $\alpha = -$

Below we give the derivations for tones 2 and 5:

		E1	E2			
			Inland	Coastal		
tone 2:	53	[+ Upper] \wedge H L	55	[+ Upper] H	33	[+ Upper] L
tone 5:	13	[- Upper] \wedge L H	33	[- Upper] H	21	[- Upper] L

There is one remaining detail to account for in the unstopped tones. Tone 3, [21], not only changes register, but also becomes a phonologically falling tone [53]. We shall therefore need a late rule of dissimilation whose effect is similar to Mandarin tone sandhi and Shanghai H-insertion.

(3) Dissimilation

$$L \longrightarrow H \quad / \quad \begin{array}{c} \wedge \\ \text{---} \end{array} L$$

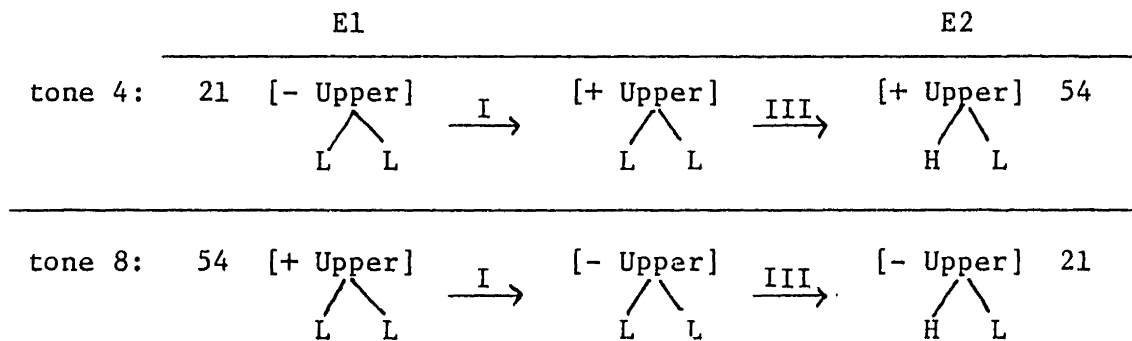
This rule as stated will apply to both tones 3 and 7:

		E1	E2			
tone 3:	21	[- Upper] \wedge L L	$\xrightarrow{\text{I}}$	[+ Upper] \wedge L L	$\xrightarrow{\text{III}}$	[+ Upper] 53 \wedge H L
tone 7:	33	[+ Upper] \wedge L L	$\xrightarrow{\text{I}}$	[- Upper] \wedge L L	$\xrightarrow{\text{III}}$	[- Upper] 21 \wedge H L

Notice that we have used some freedom here in deciding whether tones reported as [21] are phonologically falling or level; taking the output

for tone 7 to be phonologically falling (and therefore is closer accord with Cheng [31] representation) allows a simpler statement of rule III, but should evidence be found to show that this is wrong, the rule could be restricted to apply only to [+ Upper] register tones.

It remains to account for the stopped tones 4 and 8. These behave in exactly the same way as 3 and 7 respectively, and merge with them if the stop is lost.



The obvious problem here is that tone 8 in E1 is phonetically [54], which is higher than would be predicted given the phonological representation proposed above, [+ Upper, L]. However, it is quite clear from the dialects in which this tone has started to merge with the mid level tone 7 that the two are tonologically identical. Speakers of this dialect may have two pronunciations for a single morpheme, one a stopped [54] and the other an unstopped [33]. The sandhi form of both is [21], with or without stop, whereas conversely the stopped and unstopped citation [21] (tones 3 and 4) have the tonologically identical but phonetically distinct stopped [54] and unstopped [33] sandhi forms respectively. We assume, then, that there is a late phonetic rule raising the stopped cases to [54].

We have now accounted for all the alternations between E1 and E2 by means of three rules, one of which also accounts for the difference between the Inland and Coastal dialects. The forms in E3 were accounted for in 1.1.3 as the result of the addition of a floating tone, [+ Upper, H], which attaches leftwards and results in a high rising or high level tone. The forms in E4 were accounted for in 1.4.3 and 4.2.1 as the result of straightforward Tone deletion and subsequent spreading.

It might be instructive to compare this analysis with that proposed by R. L. Cheng (1973) using Wang's (1967) features. Starting with the simplest case, he formulates the rule for the forms in E4 as follows:

$$(4) \quad [+ \text{ Falling}] \longrightarrow [- \text{ Falling}] / \text{ ______ } a^{53}$$

Although this rule is very simple, it lacks explanatory power, since it requires a morphological rather than a phonological environment. Cheng follows it with a functional explanation for the inability of the preceding tone to fall when in such close juncture with the following affix, but it is not clear why a sequence such as [53 53] should be any lengthier than [33 53], since both involve a transition from mid to high level at the syllable boundary. Notice, on the other hand, that the analysis proposed here predicts that the output will be H, since it is the result of spreading from the affix, which itself begins with a H tone.

Cheng's rule for the first syllable of reduplicated adjectives is less attractive. He writes:

$$(5) \quad [- \text{ high}] \longrightarrow \left[\begin{array}{l} + \text{ rise} \\ - \text{ short} \\ + \text{ high} \end{array} \right]$$

E1	2 53	3 & 4 21(?)	7 & 8 33 (54?)	5 13	1 55
	+ u ∧ H L	- u ∧ L L	+ u ∧ L L	- u ∧ L H	+ u ∧ H H
(1) Register Switch ɔu → ɔ̄u / <u> </u> ∧ ɾT ɾT		+ u ∧ L L	- u ∧ L L		- u ∧ H H
(2) Dialectal Recessive Deletion - ɔH — ∅ // ɔ H Inland: ɔ = + Coastal: ɔ = -	I: + u H C: + u L			I: - u H C: - u L	
(3) Dissimilation L → H / <u> </u> L ∧		+ u ∧ H L	- u ∧ H L		
E2 Inland	55	53 (54?)	21(?)	33	33
Coastal	33	same	same	21	same

E3 & E4

	2 55 (C:33)	3, 4 53 (54?)	7, 8, 5, 1 21(?), 33(C:21), 33
E2	$\begin{array}{c} +u \\ \\ H \end{array} \left(\begin{array}{c} +u \\ \\ L \end{array} \right) \text{ (C:)} \quad \begin{array}{c} +u \\ \wedge \\ H \ L \end{array}$	$\begin{array}{c} +u \\ \wedge \\ H \ L \end{array}$	$\begin{array}{c} -u \\ \wedge \\ H \ L \end{array} \quad \begin{array}{c} -u \\ \\ H \end{array} \left(\begin{array}{c} -u \\ \\ L \end{array} \right) \quad \begin{array}{c} -u \\ \wedge \\ H \ H \end{array}$
$\begin{array}{c} +u \\ \\ H \text{ insertion in E3} \end{array} \quad \begin{array}{c} -u \\ \\ H \end{array}$ E3			$\begin{array}{c} -u \\ \wedge \\ (T_1) \ T_2 \end{array} \quad \begin{array}{c} +u \\ \\ H \end{array}$ 35
T deletion in E4 E4	$\begin{array}{c} +u \\ \\ \end{array} \quad \begin{array}{c} +u \\ \wedge \\ H \ L \end{array}$ 55 53	$\begin{array}{c} +u \\ \\ \end{array} \quad \begin{array}{c} +u \\ \wedge \\ H \ L \end{array}$ 55 53	$\begin{array}{c} -u \\ \\ \end{array} \quad \begin{array}{c} +u \\ \wedge \\ H \ L \end{array}$ 33 53

This rule shares one property with our analysis: it is restricted to applying to those tones which change phonetically in this environment. Again, however, it lacks explanatory power; we take it to be more than coincidence that in many of the dialects reduplication processes are associated with floating high tones (for example, see Mandarin, 1.1.2 and Cantonese, 1.1.1). Amoy differs from other dialects in that the high falling tone does not become high level in this environment, and this has resulted in a reanalysis under which the high tone is only added to [- Upper] register tones. Under Cheng's analysis there is no immediately obvious way of capturing the similarities between such processes.

Lastly, let us consider how Cheng's analysis would account for the Coastal dialect. The relevant changes from E1 to E2 are the result of the following two rules (R. L. Cheng p. 9):

$$(6) \quad \begin{bmatrix} \alpha \text{ high} \\ \beta \text{ falling} \end{bmatrix} \longrightarrow \begin{bmatrix} \beta \text{ high} \\ -\alpha \text{ falling} \end{bmatrix} / \begin{bmatrix} \text{---} \\ -\text{short} \\ -\text{rising} \end{bmatrix}^4$$

$$(7) \quad [+ \text{ rising}] \longrightarrow [- \text{ rising}]$$

These apply in the Inland dialect with which Cheng is concerned as follows:

$$\text{tone 2:} \quad \begin{array}{ccc} & 53 & \\ & \begin{bmatrix} + \text{ high} \\ + \text{ falling} \end{bmatrix} & \xrightarrow{1} & \begin{array}{c} 55 \\ | - \text{ falling} | \end{array} \end{array}$$

$$\text{tone 5:} \quad \begin{array}{ccc} & 13 & \\ & \begin{bmatrix} - \text{ high} \\ + \text{ rise} \end{bmatrix} & \xrightarrow{2} & \begin{array}{c} 33 \\ [- \text{ rise}] \end{array} \end{array}$$

The two rules will never apply to the same forms, so no sort of re-ordering will explain the Coastal dialect. Cheng points out that a dialect which has the lower form of tone 5, like the Coastal dialect, could be accounted for by combining rules 1 and 2 as follows:

$$(8) \quad \begin{bmatrix} \alpha \text{ high} \\ \beta \text{ falling} \end{bmatrix} \longrightarrow \begin{bmatrix} \beta \text{ high} \\ -\alpha \text{ falling} \\ - \text{ rising} \end{bmatrix} / \begin{bmatrix} \text{---} \\ - \text{ short} \end{bmatrix}$$

He does not discuss the dialect described by Hsieh, but presumably it could be accounted for by replacing [β high] in the output of rule 3 by [- high]. For Cheng, then, the dialect difference involves some fairly substantial differences in the form of the rules, and rules which are rather powerful and unusual in the first place, involving as they do two separate occurrences of variables linking different features. Under our analysis, the two dialects described by Hsieh differ only in the 'dominance' of H versus L, which shows up in the rule as a difference in the single value +/--. The dialect mentioned by Cheng, in which only one tone has different values, has a simple rule of final tone deletion, no matter what its value:

$$(9) \quad [\alpha H] \longrightarrow \emptyset / [-\alpha H] \text{ ---}$$

as the reader may easily satisfy himself.

It should be clear, I hope, that Cheng's solution lacks the explanatory power of an analysis that makes use of the possibilities of autosegmental phonology in general, and of Register and Tone levels in particular. It remains only to make specific certain other facts

about Amoy tonal phonology, starting with neutral tone.

We have suggested that neutral toned syllables are non-branching, toneless syllables. The fact remains that they are realized on the surface as low pitched irrespective of the tone of the preceding syllable. This suggests that there is an insertion rule that inserts the complex [- Upper, L] on such syllables, or other rules that have the same effect. Now the reader may also recall that in addition to tone being deleted in ordinary (i.e., SW neutral toned) feet, Tone (but not Register) is deleted in the special complex WS feet formed by the affix a⁵³ (see 2.1.2). Assuming that these two types of tone deletion are related facts, they can be collapsed into a single rule as follows:

(10)

$$T \longrightarrow \emptyset / \begin{array}{c} \uparrow \\ W \\ | \\ \text{---} \end{array}$$

Register is neutralized only in foot-final W positions:

(11)

$$R \longrightarrow [- \text{Upper}] / \begin{array}{c} W \\ | \\ \text{---} \end{array} \downarrow$$

The reader can compare these rules with those for Mandarin in 5.1 ((6) and (7)): formulated in this way, they are identical for both languages. Since Mandarin has no WS feet both rules will always apply in the same contexts, whereas in Amoy only the first rule applies in some cases. The last rule we need is one that inserts a L tone in

Shanghai, which also inserts [- Upper, L] in W position, although under quite different circumstances. The correlation between [+ Upper, H] and S on the one hand, and [- Upper, L] and W on the other is very striking in language after language. In non-tonal languages S position is usually associated with H pitch and in pitch accent systems rules associating the first H with the starred syllable are common, whereas rules associating the first L with the starred syllable are rare or unknown. Nothing I have said explains this correlation, but it is undoubtedly something that should fall out naturally from a fully developed tonal phonology.

The last topic in this section is the phenomenon of resyllabification and voicing found before a⁵³, and optionally before vowel-initial neutral toned syllables.⁵ The analysis below follows very closely one proposed by Kiparsky (1979) for English flapping and aspiration. To recapitulate, the facts are as follows. Resyllabification and voicing, which always go together, are found in two situations, both of which are foot-internal. The first is before a⁵³:

$$\text{khap}^{21} \text{ a}^{53} \longrightarrow \text{khab}^5 \text{ ba}^{53}$$

and the second is before some neutral toned affixes:

$$\text{tek}^{21} \text{ e} \longrightarrow \text{teg}^{21} \text{ ge}$$

Obviously Amoy, like Mandarin, has the preferred syllable structure of CV, and then foot-internally strings are resyllabified if necessary to conform to this template (see 5.1 and 2.2.2). The difference comes in

the voicing facts. In Mandarin any unaspirated syllable-initial stop in second position in a foot becomes voiced, whether underlyingly syllable-initial, or derived. In Amoy only the derived stops voice, so there is no voicing in:

the⁵⁴ tsau 'take it away'

This suggests that two processes may be involved, one of which applies when the stop is still syllable-final, and another when it becomes syllable initial. Suppose that syllable-final stops become lax:

(13) C → [+ lax] / _____]_σ

This is followed by resyllabification, and then by a rule voicing syllable initial lax stops:

(14) C → [+ voice] / [_____]_σ
 [+ lax]

Any outputs of (13) that do not become subject to (14) will eventually become glottalized:

(15) C → [+ constricted glottis]
 [+ lax]

This will result in derivations like the following:

	tek ²¹ e	the ⁵⁴ tsau
<u>(13)</u> →	tek ²¹ e [+ lax]	N/A
<u>resyll</u> →	te ²¹ ke [+ lax]	N/A
<u>(14)</u> →	te ²¹ ge	N/A

Note that resyllabification will in fact remove the stop from the coda of the first syllable rather than copying it onto the next syllable: Cheng's data imply gemination, but no instrumental data are available, and the speaker's knowledge of the morpheme structure may well account for any intuition that the consonant is still attached to the first syllable. The situation is exactly akin to English flapping (or, more accurately, tapping) (and indeed in the case of the alveolar a real flap can be heard in Amoy, often transcribed with a lateral).

This completes the analysis of the tonal phonology of Amoy.

5.4 Fuzhou

In section 4.4 I argued that the tonally conditioned vowel alternations of Fuzhou could be simply accounted for by a rule which raised vowels in the context of [+ Upper] Register. In this section I will examine the tone sandhi system of Fuzhou and show that this assumption also allows for a straightforward analysis of the complex sandhi system.

Unlike the Southern Min dialects (e.g., Amoy), the Northern Min

dialect of Fuzhou has a sandhi system in which the sandhi form of each tone varies depending on the following tone. A single tone may therefore have up to three sandhi forms in addition to its citation form (which occurs pre-pausally). The tone changes are shown in the following chart:

Second word First word	Tones which trigger vowel raising						
	44	52	<u>4</u>	22	12	<u>13</u>	242
44							
12							
242							
(13)							
44							
12							
242							
(13)							
52							
<u>4</u>							
22							
<u>13</u>							

The data comes from a variety of sources, which disagree on many points, as I shall point out. The main sources are T'ao (1930), Chao (1933), Hanyu Fangyan Cihui (1964), Wang (1969), HYFYGY (1960), Lan (1953). Obviously much of the disagreement may result from sub-dialectal differences, but some, especially in the notation of tonal contours, may rather illustrate the real problem in tonal phonology posed by different field workers' perceptions of the same facts: one man's 24 is another man's 35, or, more seriously, one man's 22 is another man's

32 and so forth. This is where instrumental work is indispensable, but still almost entirely lacking. I shall therefore be forced to 'interpret' differing notations at various points.

The above chart provides three ways of classifying tones:

- (i) Which set of vowels do they take? (Do they condition vowel raising?)
- (ii) What are their sandhi forms?
- (iii) How do they act as contexts for sandhi?

(Note that the differences in precise forms given by the various writers do not affect the ways in which the tones group in these respects, except that in some cases they suggest subdivisions within the larger groupings.)

I would now like to make a rather reasonable assumption about the character of tone rules: tone rules are local rules. This means that the context for a tone rule will always be an adjacent tone, rather than a tone at a distance. As a result, if the tone of a morpheme changes before another morpheme, the rule must be one that changes the second toneme of the first morpheme when followed by the first toneme of the second morpheme. We can then see that classification (ii) above, by sandhi form, is a classification by second toneme, whereas classification (iii), as contexts, is a classification by first toneme. I have already argued that (i), classification by vowel height, is a classification by Register, so the three together give a complete picture of the underlying tone.

Using these three criteria, the tones can be represented as follows:

		Second Toneme	
		H	L
First Toneme		44, 12, 242	52, <u>4</u> , 22, <u>13</u>
H	44, 52, <u>4</u>	[+ Upper]: 44 HH	[+ Upper]: 52, <u>4</u> HL
L	22, 12, <u>13</u> , 242	LH [- Upper]: 12, 242	LL [- Upper]: <u>13</u>

For example, [44], [52] and [4] always act in the same way as a context for sandhi. We can therefore assume they have the same first Toneme, and since [52] is a falling tone that first Toneme must be H. On the other hand, [44] differs from [52] and [4] in its sandhi forms. [52] and [4] always merge, and therefore presumably have the same second Toneme, which must be L (since [52] is falling). [44] must therefore differ in ending with a H Toneme.

One tone is not included in the above chart: the sandhi tone [35]. Since it triggers the vowel raising rule it must be [+ Upper], and since it is rising it is LH. One other detail: the tones [12] and [242] behave in every way identically in sandhi, but they must be differentiated somehow since their citation forms are different. The simplest assumption is that [242] is underlyingly [- Upper, LHL], and

an early simplification rule deletes the final L, thus merging it with 12 [- Upper, LH]). Tones 52 and 4 are tonally identical, but differentiated by the presence or absence of the glottal stop.

The complete tonal inventory is given below:

[+ Upper]		[- Upper]	
44	HH	12	LH
52, <u>4</u>	HL	<u>13</u>	LL
22	LL	242	LHL
35	LH		

Before proceeding to an analysis of the sandhi rules, there are certain representations in the above table that need some justification. In most cases the surface form is that predicted by the phonological representation, but in a few cases there is an apparent discrepancy:

(i) Tone [4]: The phonology suggests this is falling, but phonetically it is high level. We therefore need a late rule of L-tone raising applying on stopped syllables only. If instead we take it to be underlyingly high level we must accept that the sandhi rules have segmental as well as tonal contexts. There is no reason why this could not be true, but let us first explore the consequences of restricting the contexts to tonal ones.

(ii) Tone [13]: This is phonetically rising, although perhaps less than suggested by [13]: Hanyu FYCH gives [23]. Further, some syllables

with this tone, probably those which have -ʔ or -∅ instead of -k (Wang, Egerod), undergo the same sandhi as [12], suggesting a final H (although there are dialects with [11], not [12] (See Wang) which suggests L). Notice the parallelism with the other stopped tone, [4]. Both appear from their ultimate sandhi forms to end in L, but both phonetically apparently end in H. As was remarked earlier, one possibility is that the sandhi rules are sensitive to segmentals. We will make a different assumption: that the sandhi contexts are tonal only, and the segmentals affect final phonetic form. In particular, on the surface all stopped tones end H, by a very late rule which will be assumed in what follows.

The rule raises final L before glottal stops:

$$L \longrightarrow H / \begin{array}{c} V? \\ | \\ \underline{\quad} \end{array}$$

and will therefore change 13 /LL/ \longrightarrow LH and 4 /HL/ \longrightarrow HH.

(iii) Tone [22]: The literature gives this variously as [33] (on a six*point scale) or falling [21] or [31]. Why then are we justified in calling it [+ Upper]? Firstly, of course, it triggers vowel raising. Secondly, if it were [- Upper] it would of necessity be H(L), to distinguish it from the other [- Upper] tones, which would group it wrongly as a context. Given that features like Upper refer to relative notions, and that the lower level of [+ Upper] 'touches' the higher level of [- Upper], this representation of [22] as [+ Upper] is no great problem, since it begins at a level that is equal to (or

higher than) all the [- Upper] tones.

We are now ready to examine the actual sandhi processes.

Inspection of the sandhi table on page 338 makes it clear that the output of the sandhi rules is always a [+ Upper] Register tone.⁶ We therefore need a rule:

$$(1) \quad R \longrightarrow [+ \text{Upper}]$$

(The context for this and all subsequent rules includes a common specification 'when dominated by a W node' as in Amoy sandhi. It will therefore not be stated in each rule.) The ordering of this rule will be discussed below.

We also need a rule of early tone simplification that merges /242/ with /12/, since their subsequent behaviour is in all respects identical:

$$(2) \quad L \longrightarrow \emptyset \quad / \quad \begin{array}{c} \nearrow \\ \text{LH} \end{array} \underline{\quad}$$

Now consider the top left hand box of the table, which involves tones ending in a H Tone when preceding tones beginning in a H Tone. The output is always a plain high tone, [44], which suggests that in some cases the initial tone has been deleted. For example, consider the sequence /12 52/. This has the underlying form:



but it surfaces as:



Now the change to [+ Upper] has already been dealt with by rule (1), but the loss of the initial L Tone requires a further rule:

(3) $T \longrightarrow \emptyset / [\underline{\quad}]$

As written the rule will delete all initial tones irrespective of what follows. So it will not only account for the level high output tone, but also for the loss of an initial L in another set of cases.

Consider the top right hand box: when one of these same H-final tones precedes a L-initial tone, the result is a falling [52] tone. In particular, consider the sequence /12 22/, which has the following underlying form:

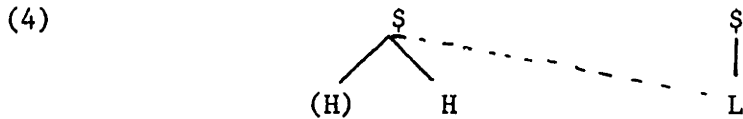


Two things must be accounted for (apart from the Register change): the lack of an initial L in the output, and the eventual fall on the first syllable. The first of these can be accounted for by rule (3), which gives:



The second is apparently the result of spreading from the second

syllable, and suggests a rule of L-spreading:



Now we must take a closer look at the formulation of rule (3), which is currently written to delete all initial tones. In addition to changing /LH/ to [H] it will therefore also change /HH/ to [H], (thereby merging /HH/ and /LH/). While this is quite acceptable, it has another less desirable effect: it will also merge /HL/ and /LL/ to [L]. It is easy to see that this is incorrect by looking at what happens to /52/ and /22/, which are /HL/ and /LL/ respectively. In all cases their sandhi forms remain distinct, so rule (3) cannot possibly be formulated in such a general form. There are two possible ways to block its occurrence in these cases. One is to restrict it to deleting initial L tones, and the other is to restrict it to applying to [- Upper] Register tones (assuming it precedes rule (1)). We will see later that the second alternative makes the right predictions in another set of cases, and therefore rule (3) must be reformulated as:

$$(3)' \quad T \longrightarrow \emptyset / \left[\begin{array}{c} - \text{Upper} \\ | \\ \hline \end{array} \right.$$

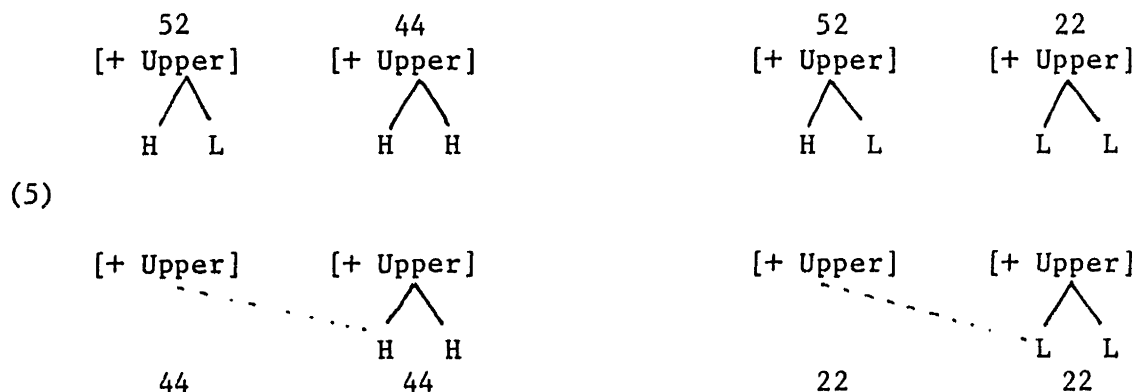
Having pinned down rule (3)' we are now able to explain why rule (4) must allow for any number of H tones. When the underlying tone is /44/, /[+ Upper HH]/, rule (3) will not apply, but rule (4), L-spreading, does apply to produce as an output the falling

[+ Upper, (H)HL] [52] tone. When the input tone is a rising /12/ [- Upper, LH], rule (3) simplifies this to [- Upper, H], and rule (4) spreads the L across to produce the falling [52] tone (after Register raising).

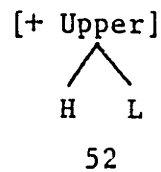
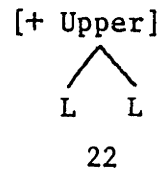
The next set of forms to be accounted for are the reflexes of /52/ and /4/, both underlyingly [+ Upper, HL]. Notice that before a tone beginning in H they surface as high level, whereas before a tone beginning in L they surface as low level. It seems then that the underlying melody HL is simply deleted in sandhi position, and the Tone of the second morpheme then spreads automatically leftwards following the WFC. Below we formulate the rule:

$$(5) \quad \text{HL} \longrightarrow \emptyset$$

which gives the following two derivations:



The last of the unstopped tones to be considered is /22/, [+ Upper, LL]. When this appears before a H tone it remains unchanged, and this is exactly what is predicted by our rules so far. Notice that none of rules (3-5) are applicable, and the sequence:

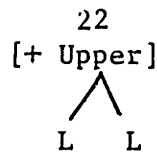
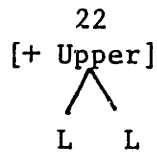


will therefore surface unchanged.

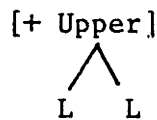
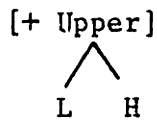
When such tones occur before a L Tone an additional complication sets in. Notice that in such contexts the output is a rising 35 tone. It is clear that the H tone is not underlying, since an underlying representation consisting entirely of L's must surface with a LH sequence. We therefore formulate:

$$(6) \quad \text{L} \longrightarrow \text{H} / \text{L} \begin{array}{c} \wedge \\ \text{---} \end{array} \text{L}$$

This will derive the right forms as follows:



(6)



Rule (6) is another example of a very common rule function in Chinese languages: the breaking up of long sequences of L Tones by dissimilation or insertion of a H Tone. Mandarin, Shanghai and Amoy all have rules of this type.

The last tones whose reflexes must be accounted for are the stopped tones 4 and 13. Under certain circumstances the final ? is

lost but in some cases it is retained.

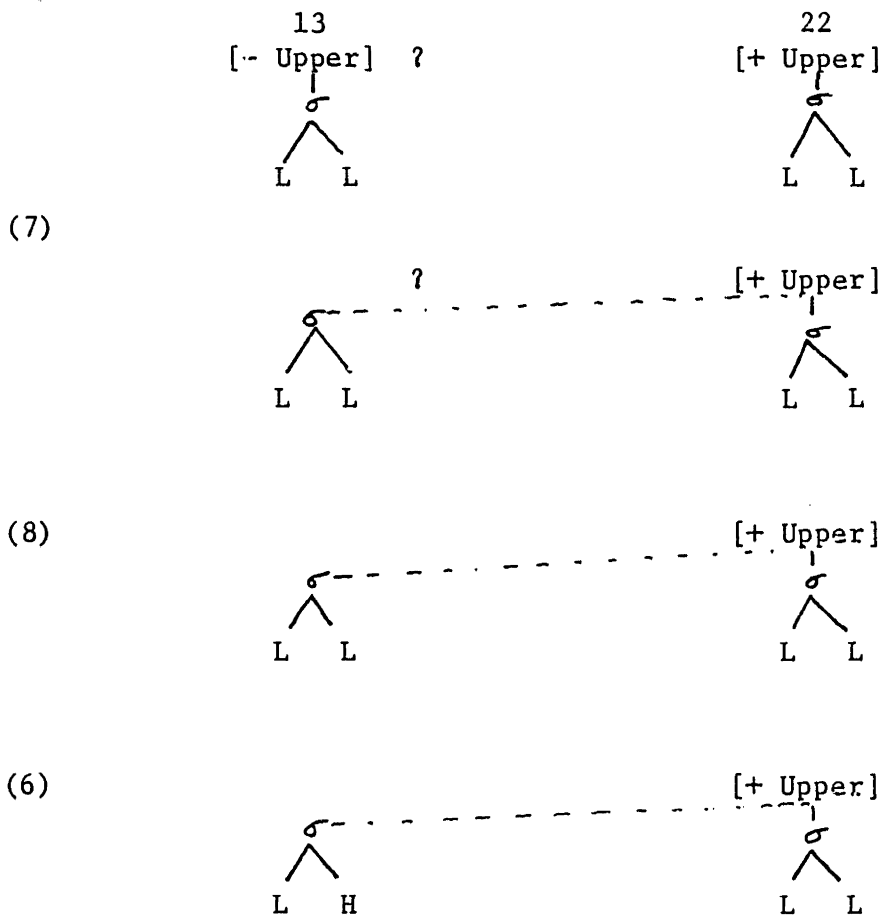
In section 4.2.1 we showed that glottal stop deletion could be stated very simply as applying to all [+ Upper] Register syllables provided that it was preceded by a rule deleting [- Upper] Register in the environment of a glottal stop. The rules can be formulated as follows:

(7) [- Upper] \longrightarrow \emptyset / _____ [+ closed glottis]

(8) ? \longrightarrow \emptyset / [+ Upper]

After rule (7) has applied to underlying /13/ tones the WFC will spread the Register of the second syllable back onto the first syllable. When this Register is [+ Upper] it will therefore create the environment for rule (8) to apply, and the glottal stop will be lost (surfacing eventually as [22] or [35]). When the Register that spreads is [- Upper] rule (8) will not apply, and the tone will eventually surface as [4].

We are now able to account for the output forms of /13/ with our existing rules. The underlying form of /13/ is [- Upper, LL]. It will therefore be subject to rule (6) dissimilating the middle L in a sequence of three. This will apply before L-initial tones, and after Register raising (rule (1)) we will correctly derive the high rising [35] that is found before [22]:



Rule (6) will also apply even when the glottal stop has not been lost (i.e., before [- Upper] Register tones). The eventual output, however, is not rising but a level [4]. Phonologically we have argued that at least some of the surface [4] tones are falling HL, which suggests that not only must the initial L be lost, but also the L of the second syllable must spread backwards. Both these processes are already known to happen in other cases, so it is only necessary to order the rules accordingly:

in glottal stop deletion) will apply, and the output will be a plain [+ Upper, LL] tone that is indistinguishable from the underlying /22/.

Below is a summary of the rules, ordered correctly. For reference the rule number as it appears in the above discussion is given to the right of each rule. In the derivations overleaf the rules are numbered in the order in which they apply.

(1) LHL Simplification

$$L \longrightarrow \emptyset / LH \text{ ______} \quad (2)$$

(2) [- Upper] Deletion

$$[- \text{ Upper}] \longrightarrow \emptyset / \text{ ______ } [+ \text{ closed glottis}] \quad (7)$$

(3) Glottal Stop Deletion

$$? \longrightarrow \emptyset / \text{ ______ } [+ \text{ Upper}] \quad (8)$$

(4) HL Deletion

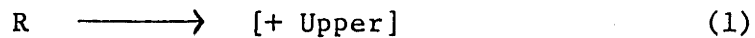
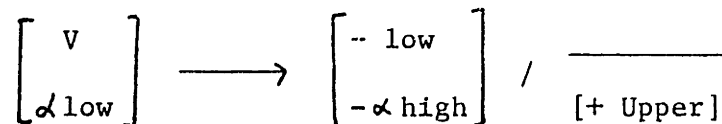
$$HL \longrightarrow \emptyset \quad (5)$$

(5) L Dissimilation

$$L \longrightarrow H / L \begin{array}{c} \wedge \\ \text{ ______ } \end{array} L \quad (6)$$

(6) T Deletion

$$T \longrightarrow \emptyset / \left[\begin{array}{c} [- \text{ Upper}] \\ | \\ \text{ ______ } \end{array} \right] \quad (3)'$$

(7) L Spreading(8) Register Raising(9) Vowel Raising

These will apply to give the right output as shown overleaf. Notice an interesting fact: with the single exception of rule (6), all the rules apply either on the register level or on the tonal level, but never on both. This argues rather strongly that the separation of the features into independent systems in this way is right; rules like (6) which make reference to both are then the marked cases, just like tonal rules which make reference to segmental contexts.

One other general comment about the rule system proposed above: with the exception of rule (5) dissimilating a low to a high tone, all the rules are rules of neutralization which delete tonemes or spread them across a wider domain, thereby reducing the eventual number of tonal contrasts. R. L. Cheng (1973) suggested that tone sandhi might profitably be viewed as conditioned by stress, with any degree of stress other than main stress reducing the number of tonal contrasts. Whether or not Fuzhou sandhi is conditioned by stress is not clear (although in Amoy it undoubtedly is), but the other part of

<u>Before L tone</u>					<u>13</u>			
44	52	<u>4</u>	22	12	Before +u	Before -u	242	
+u ^ HH	+u ^ HL	+u? ^ HL	+u ^ LL	-u ^ LH	-u? +u ^ LL	-u? -u ^ LL	-u ^ LHL	
(1) LHL Simplification							-u ^ LH	
(2) [-u] Deletion					? [+u] ^ L L	? -u ^ L L		
(3) ? Deletion					+u ^ H L	[+u] ^ L L		
(4) HL Deletion					+u ^ L	+u ^ L		
(5) L Dissimilation					+u ^ L H L	+u ^ L H	? -u ^ L H	
(6) T Deletion					-u H	? -u ^ H	-u H	
(7) L Spreading					+u ^ H H L	-u ^ H L	? -u ^ H L	-u ^ H L
(8) Register Raising					+u ^ H L	+u? ^ H L	+u ^ H L	
(9) Vowel Raising								
52	22	22	35	52	35	<u>4</u>	52	

Before H tone (always +u)

	44	52	<u>4</u>	22	12	<u>13</u>	242
	$\begin{matrix} +u \\ \wedge \\ HH \end{matrix}$	$\begin{matrix} +u \\ \wedge \\ HL \end{matrix}$	$\begin{matrix} +u? \\ \wedge \\ HL \end{matrix}$	$\begin{matrix} +u \\ \wedge \\ LL \end{matrix}$	$\begin{matrix} -u \\ \wedge \\ LH \end{matrix}$	$\begin{matrix} -u? \\ \wedge \\ LL \end{matrix}$	$\begin{matrix} -u \\ \wedge \\ LHL \end{matrix}$
(1) LHL Simplification							$\begin{matrix} -u \\ \wedge \\ LH \end{matrix}$
(2) [-u] Deletion						$\begin{matrix} ? \dots +u \\ \wedge \quad \downarrow \\ L \quad L \quad H \end{matrix}$	
(3) ? Deletion			$\begin{matrix} [+u] \\ \wedge \\ H \quad L \end{matrix}$			$\begin{matrix} \quad \quad +u \\ \wedge \quad \downarrow \\ L \quad L \quad H \end{matrix}$	
(4) HL Deletion		$\begin{matrix} [+u] \\ \dots \\ H \end{matrix}$	$\begin{matrix} [+u] \\ \dots \\ H \end{matrix}$				
(5) L Dissimilation							
(6) T Deletion					$\begin{matrix} [-u] \\ \downarrow \\ H \end{matrix}$		$\begin{matrix} [-u] \\ \downarrow \\ H \end{matrix}$
(7) L Spreading							
(8) Register Raising					$\begin{matrix} +u \\ \downarrow \\ H \end{matrix}$		$\begin{matrix} +u \\ \downarrow \\ H \end{matrix}$
(9) Vowel Raising							
	44	44	44	22	44	22	44

Cheng's thesis -- the reduction in contrasts -- is strongly supported by this analysis.

As an appendix to this section, it is of some interest to note that there is a closely related dialect, Fuqing, which exhibits the same phenomenon of vowel raising in the same historical tonal categories. The tones themselves, however, are different. It might therefore seem that a phonological solution is doomed to failure, and that the traditional Sinologist's approach which views the alternations as morphological hangovers of an earlier stage in the language is correct. If we could find a dialect in which the tones have changed and so have the resulting vowel alternations we would have strong support for a phonological analysis. Fuqing is unfortunately not such a case, but neither does it support the traditional analysis: on the contrary, it is subject to exactly the same analysis as Fuzhou for one crucial reason: although the tones are different, they have not changed register. We can therefore still retain the assumption that Fuqing has a rule raising vowels in Upper register tones. Below we give the tones of Fuqing subdivided according to whether or not they trigger vowel raising:

Vowel Raising [+ Upper]			Non-Raising [- Upper]		
/HL/	51	(44)	/LL/	11	(12)
/HH/	44	(52)	/HL/	41	(242)
/LL/	31	(22)	/H(L)/	<u>3</u>	(<u>13</u>)
/H(H)/	<u>5</u>	(<u>4</u>)			

(The values in parentheses are the equivalent tones in Fuzhou. Tones enclosed in slashes are possible underlying forms, but they cannot be accurately determined without a full analysis of the sandhi system.)

5.5 Cantonese

By comparison with the languages discussed earlier, Cantonese has a very simple tonal phonology. The most noticeable tone changes are the morphologically conditioned ones of the changed tone that have already been discussed in some detail in sections 1.1.1, 1.3.1, and 4.1.1 and 4.2.1. I concluded that the changed tone is the result of adding a complex [+ Upper, H] to the basic morpheme, and that this morpheme could either be underlying (e.g., the diminutive and vocative) or the result of deletion of segmentals (e.g., the perfective suffix).

Cantonese has no noticeable word or phrase stress (although of course contrastive stress can be placed on any morpheme). Hashimoto (1972: 101) also reports neutral stress, under which particles may lose their tone. If this is right it is another example of a very common phenomenon, the loss of tone on all but one syllable in a foot. Presumably the tone of the preceding syllable then spreads, but Hashimoto does not detail the phonetic results of neutral stress.

The lack of sandhi processes may be related to this lack of stress: notice that stress has been the conditioning factor in the sandhi systems investigated elsewhere in this section. Be that as it may, Cantonese has only one clear sandhi rule, which changes the high falling tone to high level before another tone that starts on a high level:

$$53 \longrightarrow 55 / \text{---} \left\{ \begin{array}{l} 53 \\ 55 \\ 5 \end{array} \right.$$

One way of formalizing this is as follows:

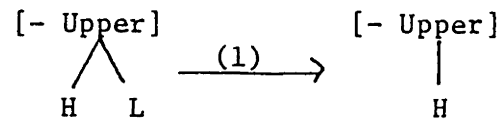
(1)

$$L \longrightarrow \emptyset / H \text{---} H$$

In this exceedingly general form the rule would apply not only to [+ Upper] Register tones, as here, but also to [- Upper] Register tones. It would thus change any falling tone to level before a tone that began on a H level. Recall the tonal inventory of Cantonese:

55 (5)	53	35
44 (4)		24
33 (3)		
22		

It appears that [53] is the only falling tone, and therefore that the second case will never arise. However, some writers including Hashimoto (1972: 112) state that the lowest tone, [22] above, actually has two variants: a basic falling /21/ that changes to [22] before another [21] (or [22]). This looks remarkably reminiscent of the tone sandhi in rule (1), but the similarity turns out to be misleading. Notice the effect of rule (1) on a [- Upper, HL] tone (presumably the underlying form of [21] if it is phonologically falling):

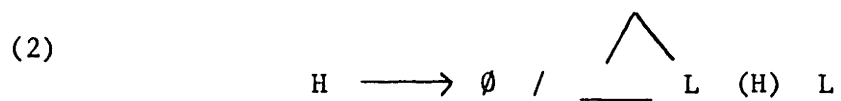


But this is very similar to the underlying form of [33] [- Upper, HH], and the rule therefore predicts a merger of /21 21/ and /33 21/.

Hashimoto is quite explicit that they remain distinct (1972: 135).

Furthermore, an alternative formulation for just this [- Upper] sandhi rule is not easy to achieve. The necessary context is apparently non-local, since it must apply before HL ([21]), LL ([22]) but not HH ([33]).

The rule would have to look like this:



There is no obvious way to collapse this with (1). At this stage one starts to suspect that something quite different might be going on, and the clue can be found in Jones (1912, as reported in Hashimoto 1972: 113). Jones points out that the falling variant normally appears in citation form or at the end of a phrase. This suggests that the fall is actually nothing more than phonetic declination, and that the level form should be taken as underlying. The sandhi rule is then an Upper register phenomenon, and, since the only falling tone [53] is Upper register the rule may be stated in the general form given in (1). There are no other tone sandhi in Cantonese, and the only interaction between segmentals and tone is the lowering effect of tense vowels in stopped syllables, for which see section 4.4.

NOTES TO CHAPTER FIVE

¹ M. Halle has suggested to me that once this rule is in the grammar it is possible to take [35] as underlyingly LL and derive the rise by this rule. If the tonal inventory were thereby simplified, this would be worth considering, but an inventory of [+ Upper] tones that includes HH, HL, and LL is not obviously simpler than one of HH, HL, and LH. Furthermore, there is no evidence whatsoever that the [35] tone is ever LL: at every stage in the derivation it behaves as LH. To postulate /LL/ as the underlying form is therefore exceedingly abstract, and does not effect any noticeable simplification of the grammar.

² The reason for the difference in environment of these rules will be clear in 5.3 on Amoy.

³ One might suggest that this [- Upper, L] tone is always inserted, but fails to surface on bisyllables. This leads to complications, since a word beginning with tone B should then surface with a falling tone over the two syllables, but in fact it is level:

*	[+ Upper]		[- Upper]
	\$		\$
	L		L

⁴ Essentially, this rule is due to Wang (1967).

⁵ Cheng (personal communication) claims it also happens before fully-

toned syllables for some affixes. Another informant finds this totally impossible, and certainly it would be very hard to explain in the framework being developed here.

⁶ When discussing the vowel alternations we noted that Wang gives cases where the output of sandhi is [12], a [- Upper] tone; these are all cases which are given in our table as [22], [+ Upper]. Wang is the only authority who gives the [12] output, and he actually differentiates two sub-cases, one of which is identical to citation [12], and another of which is a low level [11] tone found only on stopped syllables.

REFERENCES

- Anderson, S. (1974) *The Organization of Phonology*. Academic Press. New York.
- - - - - (1978) "Tone Features." In V. Fromkin, ed., *Tone: A Linguistic Survey*. Academic Press. New York. pp. 133-176.
- Asongwed, T. and L. Hyman (1976) "Morphotonology of the Ngamambo Noun." In L. Hyman, ed., *Studies in Bantu Tonology*. University of Southern California, Los Angeles.
- Bahl, K. C. (1957) "Tones in Punjabi," *Indian Linguistics* 17: 139-147.
- - - - - (1958) "A note on tones in W. Punjabi," *Indian Linguistics* 18.
- Barker, M. E. (1966) "Vietnamese-Muong Tone Correspondences." In *Studies in Comparative Austroasiatic Linguistics. Indo-Iranian Monograph 5*. Mouton: 9-25.
- Bearth, T. and H. Zemp (1967) "The Phonology of Dan(Santa)." *Journal of African Linguistics* 6: 9-29.
- Becker, D. A. (1979) "Speech error evidence for autosegmental levels." *Linguistic Inquiry* 10.1: 165-7.
- Bernot, D. (1979) "Un point de syntaxe Birmane." Paper given at the 12th International Conference on Sino-Tibetan Languages and Linguistics, Paris 1979.
- Bodman, R. W. (1978) *Poetics and Prosody in early medieval China: a study and translation of Kūkai's Bunkyō Hifuron*. Cornell University Ph.D. dissertation.

- Boyce, C. (1980) "Min sandhi in verse recitation" *Journal of Chinese Linguistics* 8.1.
- Brown, J. Marvin (1965) *From Ancient Thai to Modern Dialects*. Social Science Association Press of Thailand, Bangkok.
- - - - - (1975) "The Great Tone Split: Did it work in two opposite ways?" In J. G. Harris and J. R. Chamberlain, eds. *Studies in Tai Linguistics in honor of William J. Gedney*. Bangkok: Central Institute of English Language, Office of State Universities.
- Chang, Kun (1953) "On the tone system of Miao-Yao." *Language* 29.
- Chao, W. and M. Young (1979) "Reforestation of the Tonal Domain: a proposal for metrical structure in tonal systems." Unpublished paper, University of Massachusetts, Amherst.
- Chao, Yuan-ren (1930) "A system of tone letters." *Le Maître Phonétique* 45: 24-47.
- - - - - (1933) "Tone and Intonation in Chinese" *Academia Sinica, Bulletin of the Institute of History and Philology* IV.2: 121-134.
- - - - - (1947) *Cantonese Primer*. Harvard University Press, Cambridge, Mass.
- - - - - (1968) *A Grammar of Spoken Chinese*. University of California Press.
- Chatterji, S. K. (1940) *Indo-Aryan and Hindu*. Ahmedabad.
- Chen, L. and J. Norman (1965) *An Introduction to the Foochow dialect*. San Francisco State College and U. S. Office of Education.
- Chen, M. Y. (1979) "Metrical Structure: Evidence from Chinese Poetry." *Linguistic Inquiry* 10.3: 371-420.

- - - - - (1980) The primacy of rhythm in verse: a linguistic perspective. *Journal of Chinese Linguistics* 8.1.
- Cheng, C-C. (1973) A synchronic phonology of Mandarin Chinese. Mouton.
- Cheng, R. L. (1968) "Tone sandhi in Taiwanese." *Linguistics* 41: 19-42.
- - - - - (1973) "Some notes on tone sandhi in Taiwanese" *Linguistics* 100: 15-25.
- - - - - (1977) *Táiwān Fújiàn huà de yǔyīn jiégòují biāoyīn fǎ*.
R. L. and S. S. Cheng: Student Book Co. Ltd., Taiwan. (Phonological Structure and Romanization of Taiwanese Hokkien)
- Cheng, T. M. (1973) "The Phonology of Taishan." *Journal of Chinese Linguistics* 1.2: 256-322.
- Chomsky, N. and M. Halle (1968) *The Sound Pattern of English*. Harper and Row, New York.
- Clark, M. (1978) a. A dynamic treatment of tone with special attention to the tonal system of Igbo. *Indiana University Linguistics Club*.
- - - - - (1978) b. Representing Tone Contours dynamically.
Unpublished ms.
- Clements, G. N. (1977) "The autosegmental treatment of vowel harmony."
In V. U. Dressler and O. Pfeiffer, eds., *Phonologica* 176.
(*Innsbrucker Beiträge zur Sprachwissenschaft* vol. 19) Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck: 111-119.
- - - - - and K. Ford (1979) "Kikuyu tone shift and its synchronic consequences" *Linguistic Inquiry* 10.2: 179-210.
- - - - - (1979) "The hierarchical representation of tone features." Unpublished paper, Harvard University.

- Dave, R. (1967) "Formant analysis of clear, nasalized and murmured vowels in Gujarati." *Indian Linguistics* 28: 1-30.
- Dell, F. (1977) "The fate of the entering tone in the Chinese dialects." Paper presented at the 10th International Conference on Sino-Tibetan Languages and Linguistics, Georgetown University, Washington, D.C.
- Downer, G. B. and A. C. Graham (1963) "Tone patterns in Chinese poetry." *Bulletin of the School of Oriental and African Studies*. Vol. xxvi.1: 145-148.
- Egerød, S. (1956) *The Lungtu Dialect*. Munksgaard, Copenhagen.
- Fay, D. and A. Cutler (1977) "Malapropisms and the structure of the mental lexicon." *Linguistic Inquiry* 8.3: 505-20.
- Firth, J. R. (1957) *Papers in Linguistics*. London: Oxford University Press.
- Fischer-Jørgensen, E. (1967) "Phonetic analysis of breathy vowels in Gujarati." *Indian Linguistics* 28: 71-139.
- Fok, Chan Yuen-Yuen (1974) *A perceptual study of tones in Cantonese*. Centre of Asian Studies. University of Hong Kong.
- Fromkin, V. (1971) *The nonanomalous nature of anomalous utterances*. *Language* 47: 27-54.
- Gandour, J. and V. Fromkin (1978) "On the phonological representation of contour tones." *Linguistics of the Tibeto-Burman Area* IV.1: 73-74.
- Goldsmith, J. (1976) a. *Autosegmental Phonology*. Unpublished MIT Ph.D. dissertation.

- - - - - (1976) b. "An overview of autosegmental phonology."
Linguistic Analysis 2.1: 23-68.
- - - - - (1979) "The aims of autosegmental phonology." In
D. Dinnsen, ed., Current Approaches to Phonological Theory. Indiana
University Press.
- - - - - (1979: forthcoming) A review of Firthian prosodic analysis.
- Gray, W. (1979) "A note on resyllabification: in re Kiparsky 1979."
MIT Working Papers in Linguistics 1: 178-81.
- Halle, M. and K. Stevens (1971) "A note on laryngeal features." MIT, RLE
Quarterly Progress Report 101: 198-213.
- - - - - and J.-R. Vergnaud (1978) "Metrical Structures in Phonology."
Unpublished paper, MIT.
- - - - - and P. Kiparsky (1979) Review of Garde. Unpublished paper, MIT.
- Hanyu Fangyan Cihui (1964) (HYFYCH) Peking: Wenzhi gaige chubanshe.
- Hanyu Fangyan Gaiyao (1960) (HYFYGY) Peking: Wenzhi gaige chubanshe.
- Hanyu Fangyan Zihui (1962) (HYFYZH) Peking: Wenzhi gaige chubanshe.
- Haraguchi, S. (1975) The tone pattern of Japanese: an autosegmental
theory of Tonology. Unpublished MIT Ph.D. dissertation.
- Harris, Z. (1944) "Simultaneous components in phonology." Language: 20:
181-205.
- Hashimoto, A. O-K. Y. (1972) Studies in Yue dialects 1: Phonology of
Cantonese. Cambridge University Press.
- Haudricourt, A-G (1954) "De l'origine des tones en vietnamien." Journal
Asiatique 242: 68-82.

- Henderson, E.J.A. (1949) "Prosodies in Siamese." *Asia Major* (New Series) 1: 189-215. (Reprinted in F. R. Palmer, ed., *Prosodic Analysis*. Oxford University Press, London 1970.)
- - - - - (1967) "Grammar and tone in South East Asian languages." *Wissenschaftliche Zeitschrift der Karl-Marx-Universität, Leipzig* 16: 171-8.
- Hombert, J-M (1978) "Consonant types, vowel quality and tone." In V. Fromkin, ed., *Tone: a linguistic survey*. Academic Press, New York.
- - - - - and J. J. Ohala and W. G. Swan (1979) "Phonetic explanation for the development of tones." *Language* 55: 37-58.
- Howard, L. (1972) A directional theory of rule application in phonology. Unpublished MIT Ph.D. dissertation.
- Hsieh, Hsin-I (1976) "On the unreality of some phonological rules." *Lingua* 38: 1-19.
- Hu, Ming-yang (1959) Haiyan Tongyuan Fangyan zhong biao diao qun de yufa yiyi. *Zhongguo yuwen* 56: 372-376.
- Huang, C-T J. (1979) a. A note on the application of Mandarin tone sandhi rules. Unpublished paper, MIT.
- - - - - (1979) b. "Metrical structure of terraced-level tones." To appear in *Cahiers Linguistique d'Ottawa* 9 (Proceedings of the 10th Annual Meeting of the Northeast Linguistics Society.) Ottawa, Canada, University of Ottawa.
- Johnson, C. D. (1970) "Formal aspects of phonological description." POLA University of California, Berkeley, Report 11.

- Jones, D. and Kwing Tong Woo (1912) A Cantonese Phonetic Reader.
University of London Press.
- Jones, R. B. (1961) *Karen Linguistic studies: description, comparison and texts.* University of California Publications in Linguistics 25.
Berkeley and Los Angeles. University of California Press.
- Kao, D. L. (1971) *Structure of the syllable in Cantonese.* Mouton.,
The Hague.
- Kennedy, G. A. (1953) "Two tone patterns in Tangsic." *Language* 29:
367-373.
- Kenstowicz, M. and C. Kisseberth (1977) *Topics in Phonological Theory.*
Academic Press, New York.
- Kiparsky, P. (1979) "Metrical Structure Assignment is Cyclic." *Linguistic Inquiry* 10.3: 421-442.
- Kjellin, O. (1977) "Observations on consonant types and tone in Tibetan." *Journal of Phonetics* 5.4.
- Lan, Yaxiu (1953) *Fuzhou Yinxi. (Foochow Phonology).* *Journal of Literature, History and Philology, National Taiwan University.*
241-331.
- Lanham, L. W. (1963) "The tonemes of Xhosa: a restatement." *Studies in Linguistics* 17: 35-58.
- Leben, W. (1971) "On the segmental nature of tone in Thai." *MIT RLE Quarterly Progress Report* 101: 221-224.
- - - - - (1973) *Suprasegmental phonology.* Unpublished MIT Ph.D. dissertation.
- Lehiste, I. (1970) *Suprasegmentals.* MIT Press, Cambridge, Mass.

- Li, Fang-kuei (1949) "Tones in the riming system of the Sui language."
Word 5.3: 262-267.
- - - - - (1977) a. A handbook of comparative Tai. University Press
of Hawaii.
- - - - - (1977) b. Lecture notes from 1977 LSA Summer Institute,
Hawaii.
- Liberman, M. and A. Prince (1977) "On stress and linguistic rhythm."
Linguistic Inquiry 8.2: 249-336.
- Löffler, L. G. (1979) "A preliminary report on the Paꞑngkhua language."
Paper given at the 12th International Conference on Sino-Tibetan
Languages and Linguistics, Paris.
- Longacre, R. (1952) "Five phonemic pitch levels in Trique." Acta
Linguistica 7: 62-82.
- Lovins, J. (1971) "The tonology of Lomongo reduplication." Studies in
African Linguistics 2: 257-70.
- McCarthy, J. (1979) Formal Problems in Semitic Phonology and Morphology.
Unpublished MIT Ph.D. dissertation.
- McCawley, J. (1970) "A note on tone in Tiv conjugation." Studies in
African Linguistics 1.2: 123-129.
- McCoy, J. (1966) Szeyap data for a first approximation of Proto-
Cantonese. Unpublished Cornell University Ph.D. dissertation.
- Maddieson, I. (1970) "The inventory of features." Research Notes
(Ibadan) 3.2/3: 3-18.
- - - - - (1976) "The intrinsic pitch of vowels and tone in
Foochow," UCLA Working Papers in Phonetics 33: 191-202.

- - - - - (1979) "Universals of Tone" in Greenberg, ed., *Universals of Language: Phonology*.
- - - - - (1979) "More on the representation of contour tones." *UCLA Working Papers in Phonetics*, March 1979: 89-92.
- Matisoff, J. A. (1973) "Tonogenesis in Southeast Asia" in L. Hyman, ed., *Consonant Types and Tone*, Southern California Occasional Papers in Linguistics No. 1.
- Mei, Tsu-lin (1970) "Tones and prosody in Middle Chinese and the origin of the rising tone." *Harvard Journal of Asiatic Studies* 30: 86-100.
- Norman, J. (1973) "Tonal development in Min." *Journal of Chinese Linguistics* 1.2: 222-238.
- O'Malley, T. (1978) Paper given at Conference of Mid-Western Linguistics Association, Minneapolis. Spring 1978.
- O'Melia, T. A. (1939) *First Year Cantonese*, Hong Kong.
- Osburne, Andrea G. (1975) *A transformational analysis of tone in the verb system of Zahao (laizo) Chin.* Cornell University Ph.D. dissertation.
- - - - - (1979) "Segmental, suprasegmental, autosegmental contour tones" *Linguistic Analysis* 5.2: 183-194.
- Pandit, P. B. (1957) "Nasalization, aspiration and murmur in Gujarati," *Indian Linguistics* 17: 165-172.
- Pierrehumbert, J. (1979) (forthcoming) "Perceptions of Fundamental Frequency Declination." *Journal of the Acoustical Society of America*.
- Pike, K. L. (1948) *Tone Languages*. University of Michigan Press, Ann Arbor.

- Postal, P. (1972) "The best theory" in S. Peters, ed., *The Goals of Linguistic Theory*. Englewood Cliffs, New Jersey. Prentice-Hall.
- Pulleyblank, E. G. (1978) "The nature of the Middle Chinese tones and their development to early Mandarin." *Journal of Chinese Linguistics* 6.2: 173-203.
- Rose, P. (1979) "Tones in the Wu dialect of Zhenhai." Paper circulated at the 12th International Conference on Sino-Tibetan languages and Linguistics, Paris.
- Rumjancev, M. K. (1972) *Ton i intonacija v sovremennom Kitajskom jazyke*. Moscow: Izdatel'stvo Moskovskogo Universitets.
- Safir, K. (1979) "Metrical structure in Capanahua." *MIT Working Papers in Linguistics* 1: 95-114.
- Sherard, M. (1972) *Shanghai Phonology*. Cornell University Ph.D. dissertation.
- Stahlke, H. F. (1975) "Some problems with binary features for tone." *Proceedings of Sixth Conference on African Linguistics*. Ohio State Working Papers 20.
- Stowell, T. (1979) "Stress systems of the world, Unite." *MIT Working Papers in Linguistics* 1: 56-76.
- Strecker, D. (1979) "Higher falls more: a tonal sound change in Tai." *Computational Analyses of Asian and African languages* 11: 30-84.
- T'ao, Yu-min (1930) *Minyin yanjiu*. *Bulletin of the Institute of History and Philology* I.4: 445-70.
- Tsuji, N. (1977) *Eight Yue dialects in Guangxi Province, China, and reconstruction of proto-Yue phonology*. Cornell University Ph.D. dissertation.

- Tung, Zhao-hun (1974) "Two lower-level tone sandhi rules in one variety of South Min." *Shi da xue bao* 18: 259-266.
- Walton, A. R. (1976) *Syllable and tone in Chinese*. Cornell University Ph.D. dissertation.
- Wang, Li (1957) *Hanyu shilu xue*. Reprinted 1973 *Zhonghua shuju*.
- Wang, Tian-chang (1969) *Fuzhou hua yin yanjiu*.
- Wang, W. S-Y (1967) "The phonological features of tone." *IJAL* 33.2: 93-105.
- - - - - and Li Kung-pu (1967) "Tone 3 in Pekinese." *Journal of Speech and Hearing Research*, vol. 10: 629-36.
- Westley, D. (1971) "The Tepepetotutla Chinantec stressed syllable." *IJAL* 37: 60-63.
- Whitaker, K.P.K. (1955-6) "A study on the modified tones in spoken Cantonese." *Asian Major* 5: 9-36 and 6: 184-207.
- Williams, E. (1976) "Underlying tone in Margi and Igbo." *Linguistic Inquiry* 7.3: 463-484.
- Woo, N. (1969) *Prosody and phonology*. MIT Ph.D. dissertation.
- Yip, M. (1980) "The metrical structure of regulated verse." *Journal of Chinese Linguistics* 8.1.
- Yu, Shou-zhen (1965) *Tang shi san bai shou xiang xi*. *Tai Ping shuju*. Hong Kong.
- Yue, Oi-kan (1976) "Interplay of vocalic segments and tones in the Yüeh dialects." In *Genetic relationship, diffusion and typological similarities of East and South-East Asian languages: Papers for the first Japan-US Joint Seminar on East and South-east Asian Linguistics: 47-59*. The Japan Society for the Promotion of Science,

Tokyo.

Zee, E. and I. Maddieson (1979) "Tones and tone sandhi in Shanghai: phonetic evidence and phonological analysis." UCLA Working Papers in Phonetics, March 1979: 93-129.

Zhang, Sheng-yu (1979) Chaoyang Fangyande Chongdie shi (Reduplication in the Chaoyang dialect). Zhongguo Yuwen, March 1979: 106-114.