PROSODY AND PHONOLOGY

by

NANCY HELEN WOO

B.S., Massachusetts Institute of Technology (1965)

M.S., Massachusetts Institute of Technology (1966)

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September, 1969

Signature of Author ________________________
Department of Modern Languages i.e. FOREIGN LITERATURE + LINGUISTICS
August 25, 1969

Certified by ________________________
Thesis Supervisor

Accepted by ________________________
Chairman, Departmental Committee
on Graduate Students

PROSODY AND PHONOLOGY

by

Nancy Helen Woo

submitted to the Department of Modern Languages and Linguistics on August 25, 1969, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

The aim of the present study is to establish a framework of distinctive features for tonal phenomena. Since a large number of languages in the world appear to be tone languages, in some sense or another, such a framework is obviously necessary in order to be able to present a complete phonological description of these languages. Before a set of features can be established, however, there is one problem which must be resolved, which is, does stress exist as a phonological feature distinct from the tone features? Although there has been no specific discussion of this problem, earlier attempts at characterizing a tone language have led some to think not. With a more detailed view of various types of tone languages, however, it is seen that stress must, in fact, exist as a separate feature from the tone features.

Perhaps the most interesting aspect of the problem of establishing a set of features for tone is the determination of the nature of these features and of the segment or segment clusters to which they are assigned. Phonological and acoustic evidence show that a segmental analysis of tone is both motivated and justified. Thus, the distinctive features of tone are, in a real sense, an extension of the present distinctive feature system as presented by Chomsky and Halle. Examples are given of the use of these features in languages where, traditionally, the tone contours have not been considered to be describable in terms of segmental features. The articulatory correlates of the features are discussed, as well as the nature of other linguistic factors which can interact with pitch.

Thesis Supervisor: G. Hubert Matthews
Title: Professor of Modern Languages

DEDICATION

I dedicate this thesis to

Morris Halle

and

Hugh Matthews

who have made me realize
the infinite wisdom behind
the saying:

焚書坑儒

ACKNOWLEDGEMENTS

During my long career as a student at M.I.T., I have had the good fortune to encounter many teachers and fellow students who have expressed their faith in my ability and who have tried to inspire me to do something with it. Although it is not altogether evident that they have succeeded, I would like to thank them now for their valiant efforts.

Although they have never had the misfortune to be my teachers, I would like to thank Bryce Leggett and Bob Weatherall for their many hours of non-linguistics discussions. They have helped me to preserve my sanity, and to keep a reasonable perspective on the world.

All the members of the Linguistics department have been very helpful to me in working out this thesis. I would particularly like to thank Ken Hale and LaRaw Maran for their help, without which I could not have written this thesis. They have not only been the source of much information and many ideas, but they have patiently suffered through listening to, and/or reading all of the early formulations of the ideas presented here. Many of the ideas in chapter IV must be attributed to LaRaw Maran. He has spent many hours working them out with me, and furnishing the data in that chapter.

I wish also to thank Ken Stevens for taking the time and trouble to teach me about laboratory phonetics. His comments and criticisms have proved invaluable.

Last, but not least, I wish to thank Morris Halle and Hugh Matthews. What I owe to these two men cannot be expressed. Their constant encouragement ("Get that thesis written, or else..."), and their criticisms and comments have provided the major impetus in getting this thesis written. I have probably learned more linguistics from long discussions and arguments with them than I have from sitting in any class.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I: WHAT IS A TONE LANGUAGE?

1.0 The question which has been used as the title of this chapter has often been asked by linguists, in particular, those working with non-Indo-European languages. The interest in the question arises from the desire to determine the particular nature of the linguistic phenomenon which is to be described in terms of pitch specifications on the sonorant segments of morphemes or words. The various definitions are of interest here because if one wishes to expand the theory of generative phonology to incorporate prosodic phenomena, and if one therefore attempts to establish a set of distinctive features for tone, as we shall be doing, then it is necessary to know which tonal phenomena are to be described using these features. In particular, it is necessary to determine if the distinctive features for tones describe different articulatory events than does the feature [stress] . Notice that this problem is independent of the system of tone features used, for the question is more basic, namely, does pitch behave independently from stress? We shall use the feature [stress] tentatively to mean acoustic "loudness", characterized by subglottal pressure. Some of the definitions which have been proposed for tone languages will be reviewed here in order to compare them and to outline the motivation behind each of the definitions. The most recently proposed definition is also the most restrictive. We will examine it in detail, for implicit in this definition is the claim that there are a number of languages in which stress or pitch is the language particular interpretation of a single more abstract feature, [accent] . This situation is analogous to the one once thought to exist with rounding and pharyngealization, which were considered to be language particular manifestations of the feature [flat] .[1] Such a claim, of course, implies that stress and pitch are mutually exclusive as distinctive features in some languages. If this is true, then under what conditions, if any, are stress and tone distinctive? That is, are there any situations justifying the existence of separate features for stress and pitch? We shall discuss these questions later. Let us first consider some of the earlier definitions of a tone language.

<u>1.1.0</u> Perhaps the most general definition of a tone language is the one proposed by D. M. Beach,[2] who calls all languages tone languages, since intonation in terms of pitch modulation is inherent in the articulation of every language. Beach says that any language which uses pitch contrastively in any manner is a tone language. This definition has not appealed to many, for

it implies that English, for example, may be considered as a tone language on the basis of such pairs as /Yes./ and /Yes?/, or even more dramatically, on the basis of /cóntrast/ (noun) and /contrást/ (verb), where the accent mark indicates primary stress, since in English, stress and high tone generally occur simultaneously.

1.1.1 A more restricted notion of tone language is given by C. M. Doke:[3]

> "Tone or intonation is the musical modulation of the voice in speech, and as such is an integral part of speech enunciation in every language. All languages are tone languages, for no one speaks in an absolute monotone. Nevertheless we do, in classification, distinguish certain languages as 'tone languages', in which the tone or tone-sequence employed is significant of meaning---semantic. In the broad classification there are two categories of tone: (i) characteristic tone and (ii) significant tone.
>
> Characteristic tone. By this I mean the particular method of grouping or succession of musical pitches which characterizes a particular language, language group or language family. In this connexion, in English, the regular beat of stress is perhaps more characteristic of the language than is the system of pitch-sequence. Nevertheless, English has a well-defined system of tone, in which there is generally a rise in pitch on emphasized (stressed) syllables and a fall at the end of sentences.......Characteristic tone, then, is the label of the language, language-group or even dialect, but is not an essential to the grammatical significance of any language.
>
> Significant tone, on the other hand, plays an active part in the grammatical significance of the language, may be a means of distinguishing words of different meanings otherwise phonetically alike and may be used to convey varying emotions.
>
> The classic language using semantic tone is Chinese in almost all its forms.........

> Grammatical tone is found in the Sudanic languages and in Bantu. A tone inflexion is sometimes the only factor governing the change in grammatical significance.....
>
> Emotional tone is much more varied and much more difficult to classify. All the emotions of interrogation, sarcasm, emphasis, surprise, doubt, irritation, exultation, command, anger, etc., may be conveyed by a modification of the tone...."

Notice here that although Doke limits tone languages to those with "significant tone", this category includes semantic, grammatical, and emotional tone. Because of this, the definition has also not been widely accepted.

1.1.2 The generally accepted definition is the one which defines a tone language as one which has lexically contrastive tone. Among those who have proposed it is Kenneth Pike.[4] From the point of view of generative phonology, this would mean a language in which the tonal phenomena are in no ways predictable, and hence must be specified in the lexicon for each morpheme. Thus, although /cóntrast/ and /contrást/ differ tonally, because the stress of both words is predictable on the basis of their grammatical categories, noun and verb, English is not considered as a tone language.

1.1.3 Recently, an attempt to define a tone language even more narrowly has been made by J. McCawley in his LSA talk, "What is a Tone Language?"[5] McCawley feels that the definition of a tone language as a language with lexically specified pitch is too broad, for it would include such languages as Japanese and Serbo-Croatian. Accordingly, he wishes to limit the definition so as to exclude Japanese, but to include, say, Mandarin Chinese. The type of language which McCawley wishes to separate from the set of tone languages are those which have sometimes been referred to as "accent" or "pitch-accent" languages. In these languages, it appears that if one is told the location of the "accent" of a word or morpheme, one can then, by applying rules, specify the total phonetic pitch contour of the word or morpheme. Thus he suggests that the proper solution to this problem is not to specify any pitch value in the lexicon for Japanese, for example, but rather to use morpheme or stem features which state the location of the "accent" for each lexical entry. McCawley then discusses four types of languages; (1) bound accent, where the tonal configuration of words is totally predictable, such as in English, (2) partially free accent, where the tonal behavior of morphemes can be characterized as being of one of two classes, as, for example, in Serbo-Croatian and Bambara, (3) free accent, where the tonal behavior of each morpheme is is determined by the specification of the location of the accent, as in Japanese, and (4) tone, where each syllable of a lexical entry is specified for pitch on its sonorants, as in the Chinese dialects. McCawley claims that the first three constitute not only a typologically different type of language, but that this distinction is linguistically significant. Thus he says:[6]

> "Would this classification be of any greater typological value than that which lumps classes 1, 2, and 3 together and divides languages into tone languages and non-tone languages? I maintain that it would not. One thing which casts doubt on the value of the 4-way classification is the fact that it is often quite difficult to tell whether what one is calling a free accent language may only require one binary accentual distinction to be marked per morpheme, or whether what one is calling a partially free accent language may really require no suprasegmental information at all to be marked in the dictionary. I have heard all three claims made for English. The reason for the difficulty in telling which is which is that outside of the dictionary entries there is no particular way in which one could expect languages of these three types to be systematically different. In the third case the morphemes come with an accent marked somewhere in them; in the first two cases some rule has to put in an accent on some syllable. However, once that has been done, exactly the same kinds of rules could apply regardless of whether the language was of type 1, 2, or 3, namely rules of the form of the English stress reduction or Japanese accent reduction, that is, rules which accent a syllable while in some sense weakening the other marked syllables.
>
> However, in a language such as Chinese which fits my definition of a tone language, the suprasegmental feature specifications of the dictionary entries are not subject to rules of that form, but are subject only to rules such as assimilations and dissimilations which are of exactly the same formal nature as the rules affecting purely segmental features such as voicing or nasality.
>
> Thus the distinction between types 1, 2, and 3 seems to be of minor typological importance, since it only affects one part of the total grammar,

> whereas the distinction between type 4 and the other types appears to have implications regarding the shape of a great many things: the dictionary, the phonological rules, and the phonetic shape of the output of the rules."

1.2.0 What must now be determined is the validity of the separation of type 4 languages from types 1, 2, and 3 languages on the grounds given, and the validity of the grouping of types 1, 2, and 3. Specifically, we wish to determine the exact nature of the difference between languages of types 2 and 3, and those of type 4, since all of these languages have been traditionally regarded as tone languages. In addition, we wish to determine if there is reason to believe that types 2 and 3 languages are indeed similar to type 1 languages. Recall that by making the claim that languages of types 2 and 3 are like those of type 1, the implicit claim is made that stress or pitch is the language specific manifestation of an abstract feature [accent] . What we shall attempt to show is that this latter claim is not true, and that stress and pitch are two separate sets of features sometimes related in a given language. The nature of the relationship between [+stress] and [+high tone] will be seen to be similar to the relationship of palatalized consonants to front vowels. That is, palatalization of consonants generally occurs when the consonant is in the immediate environment of a [-back] vowel. Similarly, [+stress] will often only be assigned to sonorants which are also [+high tone] , particularly if the language only makes a distinction between + and - [high tone] . This will prove to be the "expected" or "unmarked" case, just as palatalization of consonants before front vowels is the expected case.

In order to show that stress and pitch are not simply different manifestations of the same underlying feature, not even for the accent languages, we must first determine the validity of McCawley's claim that the distinction between type 4 languages and the other types "appears to have implications regarding the shape of a great many things: the dictionary, the phonological rules, and the phonetic shape of the output of the rules."[7] If indeed there is a difference in all three areas, lexicon, rules, and phonetic output, it would then appear that the notion of some abstract feature [accent] would be justified, for it would behave so differently from the pitch features of the type 4 languages, the true tone languages. Thus it would follow that stress and pitch could not be distinctive in any language other than a type 4 language, if even there. Let us now consider in detail what were thought to be the implications regarding the phonological rules.

The primary reason for McCawley's belief that

Japanese should be considered phonologically similar to English in its prosodic system comes from the existence of the phenomenon in Japanese which he calls "accent reduction". In any given phonological word in Standard Japanese, there can be at most one syllable which is phonetically high pitched and followed by a low pitched syllable. Such syllables have been traditionally termed the accented syllables. There are, of course, words in which no syllable is accented and, hence, in which there is no sequence of a high toned syllable followed by a low toned syllable. An example of such an item is /miyako/ <u>city</u>, which has the phonetic pitch configuration /miyákó/.[8] (The acute accent will be used to indicate high tone and no mark will indicate non-high or "low" tone.) If a lexical item is accented on the last syllable, the phonetic tone configuration will not be different for the isolated word than that of an unaccented word. Thus, /atama̋/ <u>head</u>, is accented on the last syllable (where " indicates an accented syllable). Its phonetic pitch contour is /atámá/, exactly the same as in /miyákó/. One can determine whether an item is accented or unaccented by observing if it is subject to "accent reduction".

(1). Accent Reduction: When a word contains more than one accented item, then all but the first accent are deleted.

For example, in the phrase /miyako de̋su/ *it's a city*, the only accent in this word occurs in /de̋su/ and therefore it remains. However, in /atama̋ de̋su/ *it's a head*, both items in the word are accented. The accent reduction rule thus applies, and the accent on /de̋su/ is deleted. An interpretive rule then provides the phonetic shape of these words. It states that:

(2). Interpretive rule: Excepting the first syllable, which is low toned unless accented, all syllables preceding and including the accented syllable are high toned, and all syllables following the accented syllable are low toned.

Applying the interpretive rule to the phonological words /miyako de̋su/ and /atama̋ desu/, one gets /miyákó désu/ and /atámá desu/. Thus one sees that in phrases of this type, the accentual difference between the unaccented /miyako/ and the last syllable accented /atama̋/ shows up phonetically. In these two examples, /de̋su/ is a clitic, and thus the phrases are phonological words. The accent reduction rule applies to delete all accents but the first irrespective of the number of accents present. Thus the phrase /atama dattara/ *if it were a head*, consists of the three accented lexical items, /atama̋/,

/da̋r/, and /ta̋ru/, where the latter two are clitics. The derivation of /atama̋ + da̋r +ta̋ru/ to yield the phonetic form /atámá dattara/ proceeds as follows.

(3). /atama̋+da̋r+ta̋ru/

(4). /atama̋+dar+taru/ (by rule (1))

(5). /atámá+dar+taru/ (by rule (2))

Other rules will apply to /dar+taru/ to give /dattara/.

McCawley feels that accent reduction (or deletion) within a word is very much like English stress reduction with the one difference that where English stress reduction states that when a vowel is assigned [1 stress] in a word, all other stressed syllables are reduced by 1, Japanese states that all accents occurring on syllables preceded by an accented syllable within a word are deleted. There are certain regular exceptions to this latter, but we shall not discuss them here. Thus the main difference between Japanese and English is said to be the fact that there is no analog to the non-primary stresses of English in Japanese within word boundaries. However, across word boundaries, there is evidence of a phenomenon which may be likened to secondary stress, and this is found in the fact that in a phonological phrase of "breath group", it is only possible for the first word to realize a phonetic high pitch. All following high toned syllables are realized as lowered highs. This is partly comparable to the "downstepping" reported in many African languages, the difference being that it only applies once in a breath group in Japanese, rather than to each successive set of high tones. For example, consider the phrase /dőo itta̋ra i̋i desyoo ka/ *how would it be best to go?* with the accented syllables after the application of rule (1) as indicated. The final phonetic contour of the phrase is /doo ittara ii desyoo ka/. The kind of downstepping usually found in African languages would also affect the item /ii/ to give it a phonetic pitch which is yet lower than that of /ittara/, rather than one which is of the same height as indicated. The phonological phrase /kabutte mitara/ *if (I) were to try putting on (a hat)*, contains three accented items in its underlying form, /kabűr + te/ and the two in /mi̋ + ta̋ra/. Rule (1) applies to the word /mi̋ + ta̋ra/ to delete the accent on /ta̋ra/ and gives /mi̋tara/. The final phonetic contour is then /kabutte mitara/, where the high tone on the second word has been lowered.

The accent reduction across word boundaries and within breath groups which yields the lowered highs, and the accent deletion within word boundaries are considered by McCawley to be two parts of the same convention. Thus he says:[9]

> "Since only one accent can occur within a word, a rule will be necessary which deletes all accents other than the highest one within a word; this rule would delete the 'tertiary accent' in mitara and yield the correct form: primary accent on the second syllable of kabutte and secondary accent on the first syllable of mitara.
>
> In the preceding section I introduced a convention that when a rule inserts an accent it eliminates all other accents present. Suppose that this convention is changed to read that when an accent is inserted it reduces all accents present (in the stretch to which the rule is applying) by one degree. Since within a word all but the heaviest accent present is deleted, this change in the convention would have no effect on the rules for compounds...."

Recall that McCawley made the claim that rules or conventions such as stress and accent reduction are unique to languages of types 1, 2, and 3. He furthermore stated:[10]

> "However, in a language such as Chinese which fits my definition of a tone language, the suprasegmental feature specifications of the dictionary entries are not subject to rules of that form, but are subject only to rules such as assimilations and dissimilations which are of exactly the same formal nature as the rules affecting purely segmental features such as voicing or nasality."

We shall now examine the validity of this claim with respect to Mandarin Chinese.

1.3 In discussions of Mandarin tone, one rarely finds mention of stress except perhaps for a statement that those syllables which occur with the neutral tone are unstressed. What is not stated is whether the unstressed quality of a syllable is always predictable from the fact that it bears a "neutral" tone or whether the fact that a syllable has a "neutral" tone is always predictable from the fact that it is unstressed. Syllables in the neutral tone are much shorter in duration than stressed syllables and are articulated with a pitch contour which is predictable from the contour of the first stressed syllable preceding them. What must now be determined is whether or not a stress assignment rule exists in Mandarin which is independent of the tone configuration of the syllables to which it applies. That is, is stress to be considered a feature independent of the pitch features in Mandarin? We shall show that it must be.

What we shall now consider is not the entire problem of stress assignment, but the special case of verb phrases with directional complements. Verbs of motion such as /lai/ 來 <u>to come</u>, /chiuh/ 去 <u>to go</u>, can be attached to verbs to indicate direction.[11] Thus they may be considered as directional adverbials. One has, therefore, verb phrases such as /jinn lai/ 進 來 <u>come in</u>, /jinn chiuh/ 進 去 <u>go in</u>, /chu lai/ 出 來 <u>come out</u>, and /chu chiuh/ 出 去 <u>go out</u>. These verb phrases may themselves be used as directional complements to form more complex verb phrases. Thus one can have verb phrases such as /na chu lai/ 拿 出 來 <u>take out</u>

(in the direction of the speaker), /na chu chiuh/ 拿出去 take away, /kai jinn lai/ 開進來 drive in, /sonq jinn chiuh/ 送進去 send in.

In the preceding examples, only the first syllable of each phrase is long and stressed and articulated with a full citation tone, that is, one of the four contours which occur on syllables uttered in isolation. The non-initial syllables are all short and unstressed with a neutral tone configuration. Thus, these phrases differ from the lexical bisyllabic compound verbs /shuo huah/ 說話 to speak, and /day shanq/ 帶上 to present. In the latter examples, both syllables are stressed with the stress pattern of the compound being 2 1 . Both syllables are also articulated with full citation contours. In the verb phrases with directional adverbials, primary stress occurs on the main verb and all of the syllables in the adverbial are reduced to [-stress] . The [-stress] syllables are phonetically realized as neutral tone syllables, that is, these syllables are not articulated with the pitch contour specified in the lexicon, but with a contour which is completely determined by the pitch contour of the stressed syllable preceding. Thus, these syllables may be realized with a pitch contour which is totally different from that specified in the lexicon. For example, the verb /lai/

来 <u>to come</u> is specified in the lexicon with a rising tone. However, when it occurs as the adverbial in the phrase /jinn lai/ 進来 <u>come in</u>, it is phonetically realized with a short low level tone. When it occurs as the adverbial in the phrase /chu lai/ 出来 <u>come out</u>, it is phonetically realized with a short high level tone.

What we must now determine is whether the unstressed quality of the syllables in the adverbials arise from a stress placement rule which is independent of any other prosodic features of the syllables, or whether the assignment of [-stress] is merely a phonetic redundancy rule determined by the fact that a syllable has a neutral tone configuration. The fact that a syllable has a neutral tone configuration might be designated, for example, by a diacritic [+neutral]. If the value of the feature [stress] is redundantly predictable from the various pitch contours, then it will be the case that even for a type 4 language, stress is not independent of the pitch features. Such a result would lead one to question if indeed stress and pitch were not always simply separate manifestations of one underlying phonological feature. The contrary result, that there is a stress assignment rule which is independent of the pitch features of the syllables to which it applies, would, of course, indicate that at least for one type 4 language, stress and pitch are independent phonological features. If that should prove to be the case, then one must further determine which are the languages in which stress and pitch function independently.

Let us now see if it is possible to state that the unstressed syllables in the verb phrase examples given above are predictable from the pitch contours of the syllables. In order to accomplish this, it will be necessary to be able to distinguish in some way which syllables are in the neutral tone. One way of doing this would be to say that these syllables are lexically specified as such, either by having a tonal specification unique to neutral tone syllables, and therefore different from the specifications of the four citation tones, or by being marked with some diacritic, say, [+neutral]. For our present purposes, the two alternatives are equivalent, so we shall assume the latter. What does having lexically specified neutral tone syllables imply with respect to the directional adverbs? Note that we are not questioning whether there are _any_ lexical items specified as neutral tone syllables in some manner, but only whether the phonetically neutral tone syllables of directional adverbs are so specified. If one maintains the position that the syllables in the directional adverbs are lexical neutral tone syllables, then one must say that for some verbs of motion, there are two lexical entries. One would be the full verb which is not marked with the diacritic [+neutral] and the other, the adverb, which is specified as [+neutral]. Furthermore, for the complex adverbials such as /jinn lai/ as in /kai jinn lai/ drive in, one must state that there are bisyllabic lexical entries in which both syllables carry the diacritic [+neutral]. These bisyllabic entries would be composed of a first syllable from the set of all verbs which can co-occur with a directional adverb, and a second syllable from the set of verbs of motion. Since the second syllable of all of these complex bisyllabic adverbials is always one of the verbs of motion which has previously been stated to exist in the lexicon and specified there as [+neutral], one may wonder why the complex adverbials cannot be considered to be the result of some type of compounding rather than as lexical entries. If the two syllables of /jinn lai/, for example, were the result of some non-lexical compounding, then one would have to say that /jinn/ existed in the lexicon with the diacritic [+neutral] also. That is, exactly as for /lai/, there would have to be two lexical entries for /jinn/, one specifying it as a verb with a falling tone contour, and the other as an adverb with the diacritic [+neutral]. However, /jinn/ by itself does not act as an adverb. It can only occur as the first member of a complex adverbial. To state that /jinn/ occurs in the lexicon twice, once as a verb, and once as an adverb, would thus entail complicated statements of the restrictions on its occurrence as an adverb.

From the preceding, one sees that it is clearly possible to propose a solution by which the syllables in directional adverbials are designated as neutral toned in the lexicon. Given that they can be so specified, it then follows that one can simply say that [-stress] for these syllables is a phonetically redundant feature. However, let us now consider the consequences of accepting such a solution. If these adverbials are lexical items, then it is obvious that the lexicon will be far more complicated than if the adverbials were derived. In fact, it will then be true that the adverbials will bear no relationship to the verbs from which they are traditionally considered to be derived. Thus one must state that the fact that the complex bisyllabic adverbials are identical in every way except for the phonetic pitch contour to the verb phrases with a simple directional adverb is purely accidental. That is, one can no longer claim that the reason one finds the complex adverbial /jinn lai/ is because the verb /jinn/ can occur with the simple adverb /lai/. Similarly, the fact that there is no complex adverbial /niann lai/ can no longer be related to the fact that /niann/念 <u>to read aloud</u> cannot occur with the directional adverbial /lai/. Perhaps the strongest claim made by this lexical solution is that the directional adverbs are not transformationally derived from the full verbs and therefore that the co-occurrence restrictions between the verb and the adverb are not of the type of verb-verb constraints. There is nothing which supports this claim, and, in fact, all the evidence leads one to think that the opposite is true. Because this last claim does not appear to be true, and because acceptance of this solution entails unjustified complications in the lexicon, we shall reject this solution. Other reasons which make the solution untenable will become apparent in the following discussion.

An alternative to stating that the syllables in the directional adverbials are lexically specified as neutral tone syllables is to state that there is a rule which assigns a feature [+neutral] to these syllables. Then one could have rules interpreting the diacritic to give the proper phonetic pitch as well as to assign [-stress] to all [+neutral] syllables. Thus, for example, in the verb phrase /sonq jinn lai/, the rule would assign the diacritic [+neutral] to the syllables /jinn/ and /lai/. Subsequent rules would specify the phonetic pitch contour and stress value. Clearly this is a more feasible way of handling the directional adverbs. However, it is equally obvious that one could assign the feature [-stress] by the same rule and then simply state the environment of the neutral tone realization rules as $\frac{\quad}{\text{-stress}}$ rather than as $\frac{\quad}{\text{+neutral}}$.

If there were no other evidence of stress in Mandarin, then both formulations of the rule would seem to be equally correct, for there would be no justification for preferring the assignment of [+neutral] or the assignment of [-stress]. However, there are other facts to be considered. For example, there are reduplicated forms which are morphologically derived rather than lexical, in which the reduplicated second syllable is unstressed with a neutral tone contour. Thus one has the verb /tzoou/ 走 <u>to walk</u>, and the reduplicated form /tzoou tzoou/ 走 走 <u>to take a walk</u>, where the second syllable is unstressed. Here one has again the choice of saying that the second syllable is assigned a diacritic [+neutral] from which [-stress] is redundantly predictable, or that a stress rule assigns the syllable [-stress] directly. Even if these were the only examples of stress in Mandarin, one would tend to prefer the rule which assigns [-stress] , for the diacritic [+neutral] serves only to avoid having a phonological rule which mentions the feature [stress] .

There are, however, other examples of stress in Mandarin, which involve the assignment of [+stress]. Polysyllabic lexical noun compounds exhibit a stress pattern such that the final syllable is given primary stress, the initial syllable secondary stress, and intermediate syllables tertiary (or quaternary) stress. Thus, in the compound noun /$\overset{2}{\text{pin}}$ $\overset{3}{\text{in}}$ $\overset{3}{\text{tzyh}}$ $\overset{1}{\text{muu}}$/拼 音 字 母 <u>alphabet</u>, the stresses are as indicated.[12] The same stress pattern is found in syntactically derived noun phrases. For example, the noun phrase /nii de/你 的 <u>yours</u>, has the stress pattern 1 0, since the second syllable is a particle and is always unstressed. When /nii de/ occurs in a genitive noun phrase as in /nii de bii/你 的 筆 <u>your pen</u>, the resultant stress pattern is 2 0 1. Once again, the ultimate syllable receives primary stress, while the initial syllable receives secondary stress. How can we account for these stress patterns? Let us assume that the word boundaries of /pin in tzyh muu/ are found as

/##pin##in##tzyh##muu##/, and that those of /nii de bii/ are found as /##nii#de##bii##/. We assume that ## exist between all of the syllables in the first example, for there is no reason to consider any of the syllables as clitics. However, in /nii de bii/, the genitive marker /de/ is clearly a clitic; thus, there is no word boundary (##) between it and the pronoun /nii/. Let us further assume that the primary stresses have been assigned to every word in each compound. That is, we assume that we already have the configuration /$\overset{1}{\text{pin}}$ $\overset{1}{\text{in}}$ $\overset{1}{\text{tzyh}}$ $\overset{1}{\text{muu}}$/. We will discuss the form of this word level stress rule later. The stress rule for compounds can then be expressed as:

(6). Place primary stress on the last stressed word in a compound noun.

In order to reduce all other stresses, we will adopt the convention used in <u>Sound Pattern of English</u> (henceforth abbreviated as <u>SPE</u>):

(7) "When primary stress is placed in a certain position, then all other stresses in the string under consideration at that point are automatically weakened by one." [13]

In order to obtain the tertiary stresses on the intermediate syllables, we must add one more rule:

(8). 2stress → 3stress/CV___(CV)*C [1stress]

Notice that rule (8) is almost identical to the stress weakening rule (rule (117)) given in SPE. Rule (8) differs in that it applies to compounds and to all of the stressed syllables preceding the primary stress excepting the first one of the phrase. We now have the following derivation for /pin in tzyh muu/.

(9). /##p$\overset{1}{\text{i}}$n##$\overset{1}{\text{i}}$n##tz$\overset{1}{\text{y}}$h##m$\overset{1}{\text{u}}$u##/

(10). 2 2 2 1 (by (6) and (7))

(11). 2 3 3 1 (by (8))

Notice that the environment of (8) includes (CV)*. This is to be interpreted as an abbreviation for an infinite schemata CVCVCV... The rule thus applies simultaneously to all of the internal syllables as seen in (11).

If stress for compounds is to be handled by rule, then obviously word level stress contours must be specified before the stress rules for compounds apply. Let us return to the problem of stress assignment in verb phrases with directional adverbs. There would seem to be no point in having a rule which assigned a diacritic [+neutral] , for it would immediately have to be interpreted as [-stress]. Moreover, because of rules such as (6), which depend on [+stress] being present, one would then have to have a separate rule assigning [+stress] to all those syllables which were not specified as [+neutral] . Thus we propose that the stress contour of verb phrases with directional adverbs is determined directly by a stress assignment rule. The rule would simply state that primary stress is assigned to the main verb in phonological words with the structure $_{VP}[V \; _{Adv}[V]]$. However, since we assume that these verb phrases are syntactically derived, rather than lexical, we must also assume that an earlier rule has deleted the internal word boundaries. That is, by the time the word level stress rule applies, the internal ##s in $_{VP}[$##sonq# $_{Adv}[$#jinn# $_{Adv}[$#lai#] #]] must no longer exist. The stress rule will then assign stress to the verb phrase cyclically, beginning with the innermost brackets. Then, as in Japanese, there is a deletion convention which deletes all but the heaviest stress within a phonological word. We now have the following derivation for the stress contour in /sonq jinn lai/:

	sonq	jinn	lai	
(12).	1	1	1	(1st cycle)
(13).	1	1	2	(2nd cycle, convention (7))
(14).	1	2	3	(3rd cycle, convention (7))
(15).	1			(deletion convention)

Notice that the deletion convention applies to /sonq jinn lai/ but not to /pin in tzyh muu/, since the latter contains internal word boundaries (##), and is therefore not a phonological word. It is quite probable that both the main stress rule and the deletion convention apply to many structures other than the verb phrases discussed here. For example, the reduplicated verb phrases such as /tzoou tzoou/ mentioned before, and prepositional phrases have the same type of internal bracketing structure and show an identical pitch and stress pattern. A late phonological rule applies to the output of the stress rules, and reduces the pitch of the [-stress] syllables to the appropriate neutral tone contour.

This is by no means a complete discussion of Mandarin stress, but it clear from the preceding that the kind of stress rules which exist in Mandarin are similar in form to the accent rules in Japanese and the stress rules in English. That is, McCawley has stated that in Japanese, within phonological words, all but the heaviest accent is deleted. Thus, from his accent rules one derives /atamá̋ dattara/ from /atamá̋ dá̋r tá̋ru/ as shown in (3), (4) and (5). There is a similar deletion convention in Mandarin applying within word boundaries, which deletes all but the heaviest stress, as seen in the derivation above for /sonq jinn lai/. The Mandarin compound rule, rule (6), is similar to the accent rules in Japanese which place an accent on the first accented word in a breath group, while reducing all other accents. These latter rules, one will recall, account for the "lowered" high tone in /kabutte mitara/. The Mandarin rule differs only in that the primary stress is placed on the last stressed syllable rather than the first, giving the final phonetic stress contour of /$\overset{2}{\text{pin}}$ $\overset{3}{\text{in}}$ $\overset{3}{\text{tzyh}}$ $\overset{1}{\text{muu}}$/ as indicated. It is obvious that a stress weakening rule such as (8) must exist, for there are speakers who have an additional weakening rule,

(16). 3stress ⟶ -stress/ under certain conditions.

The effect of rule (16) is simply that the medial syllables in a compound such as /pin in tzyh muu/ are articulated as unstressed, neutral tone syllables, rather than as weakly stressed syllables.

The Mandarin stress rules not only specify the stress contours of phonological words and compounds, they also determine the environment for tone reduction. The neutral tone contours <u>only</u> occur on [-stress] syllables, and hence all rules which specify the phonetic shape of any given neutral tone syllable must have as its environment, $\underset{\text{-stress}}{\underline{\qquad\quad}}$. We will discuss the actual nature of these rules in chapters II and III. From the Mandarin discussion, one can see that the following two statements are true. First, the Mandarin phonology contains stress rules which are similar in form to both the Japanese accent rules and English stress rules. Second, the lexical pitch features of Mandarin syllables are affected by the stress rules in that tone reduction occurs whenever a syllable is reduced to [-stress] . We have not yet spoken about tone sandhi in Mandarin, but we would like to mention now that in Mandarin, as well as in other Chinese dialects, one of the conditions which must be fulfilled before a sandhi rule can apply is that the syllable must be [+stress] . Thus, only stressed syllables are subject to sandhi. Stress rules, therefore, not only have the effect of reducing tone, but also serve to determine part of the environment for tone sandhi.

The preceding invalidates McCawley's statement about the type of rules which are supposed to exist in accent languages but not in tone languages. Recall that he said that "in a language such as Chinese, the suprasegmental feature specifications are not subject to rules of that form, but are only subject to rules such as assimilation and dissimilation which are of exactly the same formal nature as the rules affecting purely segmental features such as voicing and nasality."

(by "that form" McCawley refers to rules of the type of the English stress rule.) It is difficult to see how the neutral tone which arises from stress placement can be regarded as anything other than tone reduction, and one which results only from stress reduction. The neutral tone of Mandarin can therefore be regarded as the tonal analog to the reduced vowel of English, for both are conditioned by the feature [-stress] . With respect to the statement that the prosodic features of tone languages are only subject to rules such as assimilation and dissimilation, one readily finds examples to show that this is not true. For example, in the Suchow dialect of Chinese, there is a tone change which is reported to occur whenever a syllable with a high rising tone precedes a syllable with a high falling tone. Whenever this configuration occurs, the falling tone becomes a high level tone.[14] In Amoy, the tone contour of a syllable depends on the position of the syllable in a phrase. Thus, the tone of any given syllable has one form when the syllable is phrase final and another when it is not.[15] In Mezquital Otomí, a bisyllabic stem occurring phrase finally (and thus also in isolation) occurs with a high tone on the second syllable irrespective of the tone of the first syllable.[16] None of these examples are instances of assimilation or dissimilation of any type known in segmental phonology.

The mistake which McCawley made was a natural one, for stress is rarely mentioned in descriptions of tone languages. The fact that McCawley's claim about the differences in the form of the rules used by accent languages versus tone languages is no longer valid does not imply that the more general claim, that Japanese is unlike Chinese, is also not valid. The claims made about the differences in the lexicon, and, as a result, in the phonetic contours in these two languages still hold. It has been suggested by McCawley[17] as well as by Halle and Chomsky[18] that in the lexicon of accent languages, there is a diacritic associated with lexical formatives rather than phonological features associated with a given vowel in a word. Thus the lexicon of an accent language would be similar to that of a vowel harmony language where it has also been suggested that diacritics be used. It is therefore clear that in an obvious way, the accent languages can be termed "tone harmony" languages. By the term "accent" we do not now include those languages such as English where no prosodic features of any type are specified in the lexicon.

1.4 While we do not dispute the claim that Japanese is unlike Chinese in its prosodic system, we do question the validity of the grouping of English with Japanese.

That is, can one still maintain the position that the accent in Japanese and the stress in English are the same phenomenon? If this is so, then it should follow that there is no stress phenomenon in an accent language. Notice that is only by showing that in <u>no</u> accent language does there exist a separate stress phenomenon that there could be any real linguistic significance to the claim that type 1 languages should be grouped with languages of types 2 and 3, our tone harmony languages. Since we have shown that prosodic rules can be found in a tone language, McCawley's type 4 languages, which have the same form as the English stress rules or the Japanese accent rules, and since these rules do, in fact, affect the pitch features of lexical items, it follows that the only difference of linguistic significance between Japanese and Mandarin, for example, lies in the shape of the lexicon. However, the lexicons of English and Japanese also differ, for the one has diacritics associated with lexical items which are later interpreted to give the phonetic pitch contours of words, and the other has no prosodic features of any type associated in any way with lexical items. Thus, unless one can show that pitch and stress never exist independently in languages of types 1, 2, or 3, the differences in the shape of the lexicon suggest that

English should not be grouped with the tone harmony languages.

1.4.1 With respect to the problem posed above, let us consider some examples from Bambara, one of the Mande languages of Africa.[19] There are two contrastive pitches in Bambara, high and non-high, or low. There is also downstepping, but it is always conditioned by the existence of a low tone. Bambara is not a lexical tone language, but rather a tone harmony, or accent, language. For any given lexical item, there are only two possible pitch patterns, "high" or "low". Bambara is therefore a type 2 language in McCawley's framework, since a binary feature is associated with every lexical item rather than a feature of location, as in Japanese. The phonetic contours of the "high" lexical items may vary according to the number of syllables in the item, nevertheless, there is only one "high" contour for any lexical item of a given length. Thus the two syllable "high" noun /yírí/ tree, has the tone configuration HH, while the three syllable "high" noun /mángoró/ mango, has the tone configuration HLH. The different configurations, however, only occur when the noun is not part of a noun phrase.

For noun phrases of the type [Noun + Adjective], the tone class of the noun determines the phonetic pitch contour of the entire phrase, irrespective of the tone class of the adjective or adjectives. For example, when the high tone noun /mángoró/ occurs with the low tone adjective /nyiman/ <u>good</u> (where /-man/ is an adjectival suffix), the resulting pitch contour of the noun phrase is /mángóró nyímán/ <u>the good mango</u>. Actually, the last syllable in the phrase will show a phonetic falling contour, but this is the result of determination.[20] (In the following Bambara discussion, we use the terms <u>high tone</u> and <u>low tone</u> to refer to the tone classes "high" and "low" of lexical formatives.) When the low tone noun /muso/ <u>woman</u> is modified by /nyiman/, the pitch pattern of the phrase is /muso nyímán/. In the phrase <u>the good black woman</u>, which consists of /muso/ + /fínmán/ + /nyiman/, the final phonetic pitch contour is /muso finman nyímán/. Here /fínmán/ <u>black</u> is a lexical high tone adjective. How can we account for these contours? As we have already stated, it is the tone class of the head noun (the first word), which determines the pitch contour of the noun phrase. Thus, we can state the following rule to account for the contours of phrases:

(17). If the first word of the phrase is low toned, then a high tone is placed on all syllables of the last word and the syllables of all preceding words are specified as [-high tone].

If the first word is high toned, then all subsequent words are high toned.

Clearly, Bambara is a tone harmony language, for its rules are quite similar to those of vowel harmony languages.

Bambara has also been considered to be a terrace tone language, because there is a downstepping phenomenon which is conditioned by the presence of low tones between high tones. In an utterance, each sequence of high tones after the initial high tone sequence is lower than the preceding high tone sequence if there has been an intervening low tone. If there is no intervening low tone, then the high tones are articulated at the same pitch level as the preceding highs. Thus, in the sentence /yírí fila bé yan/ *two trees are here (there are two trees here)*, the pitch pattern is /¯¯__¯_/. In the sentence, /muso bé yírí lá/ *a woman is in a tree*, the pitch pattern is /__¯¯¯¯/ where all the high tones are realized at the same pitch level since there is no intervening low tone.

What is of interest now is the fact that in addition to the purely tonal phenomena just discussed, Bambara also has rules of stress assignment. We will only discuss this with respect to high tone lexical items and noun phrases of the type described above, [Noun + Adj] .

In both situations, primary stress occurs on the last high toned syllable in the first sequence of high tones. Thus, in /mángoró/, the stress is placed on the first syllable, giving the stress contour as /m$\overset{1}{a}$ngoro/. In contrast, the primary stress is placed on the last syllable in /yírí/,/mángóró nyímán/, and /muso nyímán/, giving the stress contours /yir$\overset{1}{i}$/,/mangoro nyim$\overset{1}{a}$n/, and /muso nyim$\overset{1}{a}$n/. Because the stress must fall on a high tone syllable if there is one, it follows that the rules determining the pitch contour of simple words and noun phrases must precede the stress assignment rule. We will not discuss how the rule applies to the noun phrases, but simply state that the phonetic stress contour contains only one stressed vowel. The phonetic correlates to [+stress] have been found to be increased duration, increased amplitude, and higher pitch.[21] Thus, in /yírí/, for example, although both syllables are high pitched, the second is higher than the first. Stress must exist separately from the pitch features in Bambara, for otherwise, how can one explain the extra length, higher pitch, and increased intensity of these vowels?

From the Bambara material, we see that at least for one accent or tone harmony language, it is reasonable to assume that stress is a phonological feature which is independent of the pitch features. Although it is true that the stress will be placed on a high toned syllable if the word has a high toned syllable, this need not imply that high pitch and stress are synonymous. Since higher pitch, as well as higher intensity, are two of the manifestations of [+stress] , it would rather appear that the correspondence between stress and high tone indicates an unmarked configuration. In particular, in a language such as Bambara, where there is already a pitch distinction between high and non-high tone, one would tend to think that the more "natural" situation would be that in which the stressed syllable coincides with a high toned one. We also see from the Bambara discussion that intensity is not the only acoustic correlate of stress, as was tentatively suggested at the beginning of this chapter. Rather, there are also consistent correlates of duration and fundamental frequency. This concurs with Lieberman's findings for English, where he notes that all three, intensity, pitch, and duration, are correlates to stress.[22] In English, however, the three correlates do not necessarily appear simultaneously as they do in Bambara. Our original definition of stress must therefore be modified, for while intensity can be one of the correlates, it need not always appear as such, and there are other correlates. With this in mind, let us now see if there are other accent languages in which stress occurs separately.

1.4.2 Serbo-Croatian is a language which has at times been thought to be a lexical tone language. More recent studies, however, show that because of the regularities in the pitch contours found in words, it is more correctly considered an accent, or in our terminology, a tone harmony language. Recall that we use the term tone harmony, rather than accent, because we wish to make it explicit that the group of languages of which we are speaking include only those in which there is some diacritic associated with each lexical formative, where the diacritic is interpreted in the phonology to give the phonetic pitch contour of the formative. Thus, English is not included in the group of tone harmony languages.

In the analysis of Serbo-Croatian given in their paper, "Serbo-Croatian Accent",[23] McCawley and Browne have shown that once the location of the marked syllable is determined, one can predict the type of accent which exists on the word. Before we continue, let us discuss the nature of the Serbo-Croatian accents. Traditionally, it is said that there are two accents, rising and falling. The falling accent is restricted to occurring on the first syllable of any given word. Rising accents can occur on any syllable except the last, hence, they cannot occur on monosyllables. The phonetic realization of these accents are the following: a monosyllabic word, and hence, one with falling accent, will have a falling fundamental frequency, with the degree of the fall depending on whether the vowel is long or short. Vowel length is distinctive in both accented and unaccented syllables. In polysyllabic words, the falling accent is realized as a falling contour on the accented syllable only if the vowel is long; otherwise, the fundamental rises slightly throughout the vowel with a slight drop toward the end. The vowel of the second syllable of word with falling accent begins with a slightly lower fundamental frequency than that of the first syllable, and the fundamental drops sharply throughout the vowel. The words with a rising accent, on the other hand, are characterized by a steadily rising fundamental on the accented syllable, with the following syllable beginning very high and then falling. Both the syllable with the rising accent and the one following it are considered to be high toned. The fall in the fundamental of the syllable following the rising accent arises from one of two sources: 1) the syllable is word final, in which case a falling fundamental is normal, since unstressed word final syllables often show this kind of behavior;

2) the syllable is not word final, in which case the second syllable after the accent would be low toned, and the fall is the result of the transition from high tone to low tone.

McCawley states in his talk, "What is a Tone Language?", that the lexicon of Serbo-Croatian need only specify the location of the high pitched syllable for any given formative. If it is the first syllable, then the word has a falling accent, and if it is any syllable other than the first, the word has a rising accent. Thus, for the item /gȍdina/, which has a falling accent, the lexicon would mark the first syllable; for the item /venčávati/, which has a rising accent on the syllable indicated, the lexicon would mark the last syllable, the one <u>following</u> the traditionally accented syllable. The lexically marked syllables would then be /gódina/ and /venčāváti/ as indicated. The phonetic pitch contour would then be derived by a phonological rule which states:

(18). The syllable immediately before the mark gets a rising phonetic accent; if the syllable with the mark is the first, it gets a falling accent." [24]

Rule (18) is essentially the rule given by Browne and McCawley. We would like to suggest a reformulation. Before doing so, however, we must consider a few more details about the phonetic pitch contours of words in Serbo-Croatian. In /venčāváti/, for example, where the diacritic indicates the lexically marked syllable and not the traditional accent, we said that the marked syllable is high pitched. It is in fact the case that all of the syllables preceding the mark are higher pitched than those following, with the highest pitch being attained on the syllable preceding the mark.[25] Thus, the phonetic pitch contour of /venčāváti/ is /-ˊ\ _ /. The syllable which receives the rising accent by (18), the one immediately preceding the lexically marked syllable, not only shows a rising fundamental frequency, but is also lengthened. Syllables which are assigned the falling accent by (18) are also lengthened. That is, an accented vowel is longer than its unaccented counterpart. The long-short opposition in vowels is preserved, so that an accented long vowel is both longer than an unaccented long vowel and an accented short vowel. How does one account for the extra length? Clearly it is not a universal phonetic property of rising and falling pitch contours that they are longer than level tones, for this is not the case in Standard Thai, for example. However, as we have seen in Bambara, length can be an acoustic correlate of stress. Moreover, in Bambara raised pitch is another correlate of [+stress] . We therefore propose the following revision of the Serbo-Croatian accent rules:

(19). Place a high tone on the marked syllable and on all syllables preceding it.

(20). Place primary stress on the syllable preceding the last high tone. If there is only one high toned syllable, place primary stress on that syllable.

Rule (19) states that all syllables preceding and including the marked syllable will be higher pitched than those following. Rule (20) puts primary stress where (18) had placed rising and falling accents. We feel that the revised rules (19) and (20) more correctly the pitch phenomenon in Serbo-Croatian. If duration and raised pitch are said to be the acoustic correlates of [+stress] in Serbo-Croatian, then the extra length of the traditionally accented syllables as well as the higher pitch is accounted for directly. The rising-falling contour found in /venčāváti/ is then similar to the rising-falling contour seen in /lighthouse/ when it occurs in the noun phrase /lighthouse keeper/. There the fact that /light/ receives primary stress is indicated by a rising pitch contour. The high pitch carries over to the beginning of the second syllable, /house/, which has tertiary stress in this phrase, and then the pitch falls rapidly.[26] The pitch contour of /gódina/, which is marked on the first syllable and therefore is stressed on the first syllable, is similarly seen to be identical to that of /rébel/ (noun), where the pitch rises on the first syllable, and then falls slightly toward the end of the vowel. The second syllable of /rebel/ begins at a lower pitch than the terminal pitch of the first syllable, and falls sharply. This is exactly the contour which we have previously stated as characteristic of the Serbo-Croatian words with falling accent, or in our terminology, those where the stressed syllable coincides with the last high toned syllable. The idea that one assigns stress rather than rising or falling accent is not particularly original. McCawley himself suggested much the same treatment in his brief discussion of Serbo-Croatian in "What is a Tone Language?"[27] There he presents a rule somewhat like (20) except that also specifies the pitch contour of the stressed syllable. Thus he says:

(21). "If the first syllable is marked, it is stressed and has falling pitch; if any other syllable is marked, the preceding syllable is stressed and has rising pitch."

We feel that it is not necessary to specify the pitch contour by rule, for we think that these contours are natural phonetic consequence of the interaction between stress and pitch.[28]

1.5 We have seen from the discussion of Mandarin that the lexical tone languages can, in fact, have rules like the English stress reduction rule or the Japanese accent reduction rule. Moreover, we have seen that the pitch features are affected by these rules. We therefore conclude that McCawley's claim about the differences in phonological rules in types 1, 2, and 3 languages, the accent languages, as opposed to type 4 languages, the lexical tone languages, is incorrect. This does not imply that accent languages are not different in a linguistically significant way from lexical tone languages, for indeed they are. We have, however, questioned the validity of the grouping of English with Serbo-Croatian and Japanese. This grouping would be linguistically significant only if stress and pitch were never independent features in any of the types 1, 2, or 3 languages. From the discussion of Bambara and Serbo-Croatian, however, we must conclude that the latter statement is not true. Since one cannot state that stress and pitch in languages such as Serbo-Croatian and Bambara are identical, we therefore conclude that there is no linguistic significance in grouping English with Japanese and Bambara, for example. Rather, since the difference between Japanese and Mandarin lies principally in the shape of the lexicon, and since the lexicons of English and Japanese differ in an analogous fashion, we feel that English and other type 1 languages should not be classed as accent languages. We suggest that a three way distinction of the following type is more linguistically relevant:

A. lexical tone languages, those in which the pitch contour of a lexical formative is specified by pitch features on every vowel, *e.g.* Mandarin, Cantonese, and Igbo.

B. tone harmony languages, those in which a diacritic is associated with each lexical formative, and where the diacritic is later interpreted to give the pitch contour of the formative, *e.g.* Japanese and Bambara

C. non-tone languages, those in which the lexicon contains no prosodic features associated in any way with formatives, *e.g.* English and Northern Tepehuan.

One can incorporate lexical tone languages and tone harmony languages into one category, tone languages, but we do not see that such a classification has any significance.

What is of significance is the determination that at least in both lexical tone languages and tone harmony languages, stress can exist as a phenomenon independent of, but not unrelated to, pitch. Stress is not unrelated to pitch in that [+stress] often has the effect of raising pitch and [-stress] of lowering or reducing pitch contours. However, stress interacts in much the same manner with vowels and vowel quality. We therefore conclude that different phonological features represent pitch and stress.

CHAPTER I: FOOTNOTES

1. See, for example, Jakobson, Fant, and Halle, pg. 31.

2. Beach, pg. 84, and pg. 102.

3. Doke, pp. 215-216.

4. Pike, pg. 3.

5. McCawley, "What is a Tone Language?", pp. 8-9.

6. _ibid_.

7. _ibid_., pg. 9.

8. All Japanese examples, both of underlying and surface forms, are taken either from McCawley, "What is a Tone Language?", or McCawley, _The Accentual System of Standard Japanese_.

9. McCawley, _The Accentual System_, pg. 183.

10. McCawley, "What is a Tone Language?", pg. 9.

11. The National Romanization system of spelling is given in the slashes (//) for each character. For a detailed explanation of this romanization, see Chao, _Mandarin Primer_, chapter II, especially section 12.

12. Chao, Mandarin Primer, pg. 26, contains a brief discussion

of this problem.

13. Chomsky and Halle, pg. 16.

14. Hanyu Fangyan Cihui, pg. 8.

15. For example, see Bodman, pp. 3-4.

16. Personal communication, Ethel Wallis.

17. McCawley, The Accentual System, pg. 193.

18. Chomsky and Halle, pp. 376-377.

19. My thanks to Charles Bird for helping me with the Bambara section. For more details about Bambara tone, see Bird (1966a, chapter V), (1966b), and (1968).

20. Bird, (1966b).

21. Personal communication, Charles Bird.

22. Lieberman, chapters IV and VII.

23. Browne and McCawley, pp. 3 ff.

24. I have rephrased Browne and McCawley's rule (2) (pg. 3) which reads "accent mark" for "mark", in order to preserve clarity in the discussion.

25. My thanks to Wales Browne for this information.

26. Lieberman, pp. 148-151.

27. McCawley, "What is a Tone Language?", pg. 4.

28. My thanks to Morris Halle for pointing out the possibility of expressing the Serbo-Croatian rules as (19) and (20). The appeal of such a formulation lies in the fact that then one can precisely state the difference between the Serbo-Croatian and Russian accentual systems. Namely, Russian has (19) in its phonology, but (20) is simplified to the following:

    (20a). Place primary stress on the last high tone.

We must state here that stress in Serbo-Croatian does not seem to be correlated with intensity. This, however, only shows that the definition of stress must be considerably more complex than originally thought.

CHAPTER II: THE DISTINCTIVE FEATURES OF TONE: A REVIEW OF OTHER PROPOSALS

2.0 If one looks at descriptions of tone in different languages, one will see that there are essentially two systems by which the tones are described. The first system is the one used most frequently by people working with Oriental languages. The second is the one used by most of the people working with African languages as well as by the majority of those working with the American Indian languages. The major difference in the two systems lies in the way in which the dynamic or moving tones are described, and the assumptions which each make about the nature of the segment or segments to which the features are assigned.

In the first system, the "Orientalist" system, the dynamic tones are usually described in terms of the features [rise] and [fall]. Moreover, the tone features are generally either associated with the entire syllable, or with just the syllabic vowel. In the second system, the "Africanist" system, the dynamic tones are represented as sequences of pitch heights, each of which is associated with some sonorant segment, usually a vowel, but not necessarily a syllabic one. In the latter system, dynamic tones are thus limited to syllables containing sonorant clusters of some type. Long vowels in these languages are usually represented as geminate clusters. Hence, a syllable whose nucleus is a long vowel may have a dynamic pitch contour.

If one is interested in establishing a set of distinctive features for tone, then the features must be able to describe the word level tonal phenomena of all languages. It must, in particular, be able to accomodate both African and Oriental languages. Wang has proposed a system of features for tone in his paper, "Phonological Features of Tone". There he proposes to specify the dynamic tones by using the features [rise] and [fall] . If one looks at the raw phonetics of the various Oriental languages, in particular, the Chinese dialects, then the set of features which Wang proposes would seem to be the right features. What we wish to determine is whether or not a system of features like Wang's is the only way in which the tones of the Oriental languages can be described, and if not, whether a system such as Wang's is the more linguistically correct. With these questions in mind, let us now look at an Oriental language to see if, in fact, its tones cannot be described as simple sequences of pitch height features as the African languages are. Clearly, our first task must be to ascertain whether the non-

Oriental languages are peculiar in their restriction of dynamic tones to syllables containing sonorant clusters. If one finds that the Oriental languages do not conform to the restrictions placed on dynamic tones by the African languages, for example, it then follows that a system of features such as Wang's is the correct system for describing the tonal contours of Oriental languages.

However, if this should prove to be the case, we are then faced with the problem of determining if, indeed, only one set of features can be used to describe all tonal phenomena. That is, if it is true that in Oriental languages a rising pitch contour can occur on a single sonorant, and must be specified as such in the lexicon, whereas African languages require that a rising pitch contour occur only if the syllable contains two sonorants, then certain difficulties arise with the use of the features [rise] and [fall]. The difficulties arise because it will then be necessary to state <u>ad</u> <u>hoc</u> restrictions on the use of the feature system, whereby, for example, the + value of the feature [rise] can occur on a single sonorant segment in some languages, irrespective of whether or not another sonorant follows it, and in other languages the + value of [rise] can occur on a sonorant only if another sonorant follows it. This peculiar type of restriction implies that the feature [rise] in the two types of languages really signifies different speech events. Clearly, there are parallel difficulties with the use of the feature [fall] . Thus, one can then ask if a system of tone features for all languages has really been established, or if one is not actually using two different sets of features. It then follows that the first step in establishing a set of features for tones is to determine the relationship of tones to sonorant clusters, especially long vowels, in those languages where vowel length is normally not discussed.

<u>2.1.0</u> If one consults descriptions of tones of any of the Chinese or other tonal Asian languages, one will find that there is generally no discussion of vowel length, for it is said that in most of these languages, vowel length is not distinctive. Thus, for example, a Mandarin phonology will list the vowels, consonants, and vowel clusters (<u>i.e.</u> diphthongs), with no statement about long or short vowels. It is generally assumed that all pure vowels are normally long, and that vocalic clusters, which are diphthongs, consist of two "short" members. The tone system of Mandarin is then described as consisting of five tones, a high level tone (˥), a mid to high rising tone (˧˥),

a high to low falling tone (˥˩), a falling-rising tone (˨˩˦), and a "neutral tone". The first four tones are the citation tones, that is, the tones which are associated with monosyllabic words uttered in isolation. The neutral tone is associated with unstressed syllables which arise either from stress assignment in phonological phrases, or from polysyllabic morphemes in the lexicon with unstressed syllables. We shall discuss the lexical representation of these polysyllables in the next chapter. For the moment, we shall simply assume that there is some manner of determining which syllables are articulated in the neutral tone. Our concern now is to determine if there is any justification in associating dynamic tones in Asian languages with long vowels. Can we specify tone as a sequence of pitch heights in these languages? Moreover, what is the nature of the sequence of segments underlying the pitch specifications? Are they necessarily sequences of vowels, or does one include nasals as well?

<u>2.1.1</u> Let us now examine the situation specifically for Mandarin. We must first clarify the notion of the "long-short" opposition in vowels. Vowels in diphthongs and before nasals are normally considered short, while vowels in open syllables are long except if the syllable is unstressed. No syllable closed by a nasal has a

diphthong as its vocalic nucleus. Thus, the Mandarin syllable may be described as having the form,

(1) $C(G)V_1\begin{Bmatrix}\emptyset \\ V_2 \\ N\end{Bmatrix}$, where $V_1 \neq V_2$. $C(G)V_1$ would be the configuration for open syllables with pure vowels, and $C(G)V_1V_2$ would be the configuration for open syllables with diphthongs. Alternatively, one could posit (2) $C(G)V_1\begin{Bmatrix}V_2 \\ N\end{Bmatrix}$ as the structure. One would then state that if $V_1 \neq V_2$ in the canonical form (2), then the same co-occurrence restrictions hold for V_1 and V_2 as in the canonical form (1). The difference between the two descriptions is very slight, only lying in the description of the open syllables with pure vowels. In (1), the pure vowels are treated as a single segment, either subject to some low level phonetic rule which says that a "pure" vowel is lengthened if the syllable is stressed, or else marked as [+long] in their underlying representations. For our purposes, it is irrelevant which of these two methods of specifying the vowel as [+long] is chosen. We shall therefore simply assume that in syllables with a single vowel as its nucleus, the vowel is lexically specified as [+long]. In (2), the pure long vowels are treated as a sequence of two identical vowels, thus requiring no low level rule to lengthen them, if one assumes that geminate vowels

are always phonetically interpreted as long vowels, as seems to be the case.

In considering the two alternatives, one sees that the real question is simply whether Mandarin syllables with a pure vowel as the syllabic nucleus should be represented lexically with a single vowel specified as [+long], or whether the underlying representation of these syllables should contain a geminate vowel cluster which is phonetically interpreted by a universal convention as a long vowel. That is, are there any reasons one should prefer to specify the vowel with the feature [+long] rather than to represent it as a geminate cluster? If not, is there any reason to believe that in Mandarin, a geminate cluster is in any way distinct from a real diphthong except in the fact that the second member of the cluster is identical to the first?

2.1.2 In order to see if there might be some physical correlates of length which would help determine which of the two alternatives given is to be preferred, we recorded several sets of Mandarin syllables in isolation. The difference between each set is the tone of the syllables. Thus, in the first set, all of the syllables are articulated with a high level tone, in the second, all are rising tone syllables, and so on. In each set, the initial consonant of every syllable is the same so that the only variable is the syllabic nucleus. All the syllables recorded are possible lexical formatives. Hence, their utterance in isolation is analogous to the utterance of single words. From these recordings we made Voiceprints and measured the duration of the syllabic nucleus. All syllables are transcribed in the National Romanization system with the character in parenthesis next to the romanization.

TABLE 2.1: CITATION TONES OF MANDARIN SYLLABLES: DURATION MEASUREMENTS OF COMPONENTS IN CSEC.

	Total Syl. Duration	Initial Cons. (+glide)	Vowel	Final Nasal
I. High level tone (˥)				
ba (八)	36	2	34	
ban (斑)	38	2	21	15
bang (梆)	38	2	21	15
bei (悲)	35	2	33	
II. Rising Tone (˧˥)				
ma (麻)	42	6	36	
mai (埋)	44	8	36	
man (饅)	48	6	23	19
mang (忙)	44	6	23	15
mau (毛)	40	8	32	
mei (沒)	38	4	34	
men (門)	42	4	17	21
meng (盟)	42	6	19	17
mi (迷)	38	4	34	
mian (綿)	48	6 (+8)	20	14
min (民)	36	2	20	14
miau (苗)	42	4 (+6)	32	
ming (明)	42	6	16	20
mou (謀)	44	8	36	

TABLE 2.1

	Total Syl. Duration	Initial Cons. (+glide)	Vowel	Final Nasal
III. Falling Tone (˥˩)				
jah (詐)	31	-	31	
jay (債)	29	-	29	
jann (佔)	27	-	27 (V+N)	
janq (仗)	23	-	23 (V+N)	
jaw (照)	25	-	25	
IV. Falling-Rising Tone (˨˩˦)				
daa (打)	53	2	51	
daai (歹)	52	2	50	
daang (黨)	52	2	35	15
dao (島)	59	2	57	
deei (得)	52	2	50	
deeng (等)	53	2	37	14
dii (底)	57	2	55	
diao (屌)	57	2 (+4)	51	
diing (頂)	53	2	36	15
duu (睹)	55	2	53	
doan (短)	61	2 (+4)	40	15

In Table 2.1, one sees that within each tone group, there is a fixed duration for the entire syllable. Moreover, the duration of the syllabic nucleus appears to be a constant also, irrespective of whether it is a long vowel, a diphthong, or a vowel + nasal cluster. Thus, /ma/ in the rising tone has both the same total duration and the same vocalic duration as /mai/. The duration of syllables with different tones is not the same, but that is not relevant, for what is significant is the fact that syllables with the same pitch configuration also have the same duration for the syllabic nucleus. Thus, although the syllabic nucleus may consist of a pure vowel, a diphthong, or a vowel + nasal cluster, there are no acoustic properties of duration which distinguishes any of the three. In particular, the one member of this set of three which could be abstractly represented as a single segment with the specification [+long] is indistinguishable in duration from the other two members, both of which are necessarily abstractly represented as two segments.

The relative duration of the tones noted in this experiment correspond to the measurements which others have made in similar experiments, notably those performed by the POLA group at Ohio State.[1] In particular, the fact that the falling-rising tone has an average duration of 1.5 times the duration of the other tones is in accordance with the results noted in other experiments.[2] Because the third tone, the falling-rising tone, is peculiar in isolation, we shall not be speaking about this particular form of that tone in the immediately following discussion.

From the previously noted fact about the duration of pure vowels, we can once again approach the problem of the representation of the canonical form of the Mandarin syllable. If one were to hold the view that these vowels were to be represented as [+long] in the lexicon, then presumably, all other vowels are redundantly specified as [-long] . Specifically, the vowels which appear before nasals would be marked as [-long]. However, with regard to the vowel plus nasal clusters, one notices immediately that the vowel occupies approximately one half of the duration of the whole sonorant cluster, VN. Moreover, one notices that these final nasals have a much longer duration than initial nasals. How is this set of facts to be expressed? It is certainly not the usual case that final nasals have an absolute duration so that they occupy half the length of the articulation of the VN combination, as can be readily observed from looking at, for example, the spectrograms for English syllables with final nasals.[3]

If one were to hold the view that tones are specified with features such as those presented by Wang, and that these features are assigned to just <u>one</u> segment in a syllable, then one must reconcile the preceding acoustic findings with the phonetic interpretation of such a representation. One could say that the phonetic interpretation of a vowel which is specified, for example, as $\begin{bmatrix} \text{-long} \\ \text{+rise} \\ \text{+high} \\ \text{-mid} \\ \text{-central} \end{bmatrix}$, is not that the vowel itself actually manifests this configuration of rising pitch, but rather that a rising pitch contour is manifested over this vowel and any following sonorants. A vowel specified as $\begin{bmatrix} \text{+long} \\ \text{+rise} \\ \text{+high} \\ \text{-mid} \\ \text{-central} \end{bmatrix}$ would be given a phonetic interpretation that says that the rising contour is realized totally on this vowel. In addition, the phonetic component must specify that the duration of the following sonorant in the first case is half the prosodic contour duration. One sees from even such a minor example as this that there is something quite unsatisfactory about such a system, for it necessitates stating explicitly what would seem to be redundant information. It also fails to capture the notion that, so far as we know, no dynamic tone can occur on a short vowel distinctively.[4] Thus, one does not ever find a syllable with a short vowel and no following sonorant where the vowel can possibly be specified distinctively as [+rise] or [+fall]. The above system does not remark on this relationship, nor can it, except in the form of a redundancy rule such as the following:

$$(3).\quad \begin{bmatrix} \text{V} \\ \text{+rise} \end{bmatrix} \rightarrow \begin{cases} \text{+long}/\_\_\_\_\_\_ \begin{Bmatrix} \# \\ [\text{-son}] \end{Bmatrix} \\ \text{-long} \end{cases}$$

To have to state such a rule explicitly, however, is unsatisfactory, for it obscures the symmetry which arises from the fact that long vowels behave as if they were a series of sonorants with respect to tone, and it denies that the relationship and the symmetry are more than accidental. Stated in other words, it obscures the fact that a vowel can be specified as [+rise] <u>if and only if</u> there is a following sonorant, where one considers $\begin{bmatrix} \text{V} \\ \text{+long} \end{bmatrix}$ as the equivalent of VV.

<u>2.1.3</u> Is there any other evidence to support the analysis of Mandarin long vowels as geminate clusters? If one looks at bisyllabic words, one will find material to support this analysis. We recorded a number of pairs of two syllable expressions where the members of each

pair were differentiated solely by the stress pattern. In each of the pairs, the first member can be described as having the stress configuration, "stressed syllable, unstressed syllable", and the second as having the configuration, "secondary stressed syllable, primary stressed syllable", where secondary and primary stress are to be understood in the usual way. The tone of the syllable in the first position of both members in each pair is the same; that is, in the set /shyng lii/, the first pair, the syllable /shyng/ has a full rising tone on both occasions of its occurrence. The unstressed version of the second syllable in each of the pairs corresponds to the traditionally termed "neutral tone" syllable. We made Voiceprints of these bisyllabic words and measured 1) the total duration of each of the syllables in each word, and 2) the duration of the syllabic nucleus for each syllable (termed sonorant duration in Table 2.2). We then computed the ratio of the sonorant duration in the unstressed version of the second syllable in each pair to the duration of the corresponding stressed syllable. (A further explanation of this ratio is given after Table 2.2.)

TABLE 2.2a: MANDARIN BISYLLABIC WORD PAIRS: THEIR STRESS PATTERNS

	Syllable Pair	Stress Pattern[1]
1a.	shyng lii (行李)	S U
b.	shyng lii (行礼)	S_2 S_1
2a.	jiow shyh (就是)	S U
b.	jiow shyh (就事)	S_2 S_1
3a.	fwu chih (福气)	S U
b.	fwu chih (服气)	S_2 S_1
4a.	guh shyh (故事)	S U
b.	guh shyh (故世)	S_2 S_1
5a.	day shanq (戴上)	S U
b.	day shanq (帶上)	S_2 S_1
6a.	yaw shyh (鑰匙)	S U
b.	yaw shyh (要事)	S_2 S_1
7a.	dong shi (東西)	S U
b.	dong shi (東西)	S_2 S_1
8a.	jann juh (站住)	S U
b.	jann juh (站住)	S_2 S_1
9a.	baw chour (報酬)	S U
b.	baw chour (報仇)	S_2 S_1

TABLE 2.2b: MANDARIN BISYLLABIC WORD PAIRS: MEASUREMENTS OF SONORANT DURATION IN CSEC.

	First syllable		Second syllable		b/a[2]
	Total duration	Sonorant duration	Total duration	Sonorant Duration	
1a.[3]	42	25	17	13	3.6
b.	42	25	51	47	
2a.	25	21	23	11	1.7
b.	25	21	32	19	
3a.	32	19	17	8	2.6
b.	29	13	32	19	
4a.	21	19	21	10	2.1
b.	21	19	34	21	
5a.	23	21	17	9	1.9
b.	23	21	21	17	
6a.	25	21	21	8	2.4
b.	25	21	32	19	
7a.	25	23	21	11	2.1
b.	29	27	44	27	

TABLE 2.2b

	First syllable		Second syllable		b/a
	Total duration	Sonorant duration	Total duration	Sonorant duration	
8a.	29	25	11	8	2.2
b.	32	29	23	19	
9a.	23	21	17	13	2.1
b.	27	25	38	32	

Notes:

1). S, U, S_2, and S_1 are used to indicate stressed, unstressed, secondary stressed, and primary stressed syllables, respectively.

2). For each pair of words, the ratio b/a is the ratio of the sonorant cluster duration in the second syllable of each member after adjustments have been made to account for the possible variation in the duration of the first syllable. Thus:

$$b/a = \frac{\text{son. dur. of 2nd syl. of (b)}}{\text{son. dur. of 2nd syl. of (a)}} \times K, \text{ where}$$

$$K = \frac{\text{total duration of 1st syl. of (a)}}{\text{total duration of 1st syl. of (b)}}$$

TABLE 2.2b

Notes (cont.):

For example, in the pair (9), K = 23/27, and b/a = 32/13 x 23/27 = 2.1

3). The numbers (1a), (1b), etc. in Table 2.2b indicate that these are the measurements for the corresponding items in Table 2.2a.

If one reviews the results given in Table 2.2, one can see that the major point of differentiation between the second syllables in each pair is in the length of the vowel. This experiment is similar to the one conducted by Pavel Kratochvíl which was reported in his paper "Disyllabic Stress Patterns in Peking Dialect". In that paper, one does not find the b/a ratio as clearly defined as in our findings, but that is primarily due to the fact that Kratochvíl measured the duration of the entire syllable under stressed and unstressed conditions, thus including in the measurements such factors as the duration of the initial sibilants. In our findings, we noticed that when the syllable under consideration began with a sibilant, in both the stressed and the unstressed versions of the syllable, the sibilant had about the same duration. The only variable articulation duration was found to lie in the sonorant series beginning with the main vowel. In every pair except for the first, the b/a ratio was approximately 2.0. The first pair is exceptional, for in (1b), the second syllable is articulated in the extra-long falling-rising tone.

In traditional terminology, the unstressed syllables would be called "neutral tone" syllables, and described as having a tone configuration which consists primarily of the specification of some pitch height which is determined by the tone of the preceding syllable. Are there any acoustic correlates to this description? In Kratochvíl's work, measurements were made of the tone contours of the different types of syllables. His findings show that under primary and secondary stress, syllables consistently show the same type of contour for any given tone. In the unstressed position, however, the contours are described as staccato, or falling, and range greatly in their fundamental frequencies. Thus, the duration of the sonorant cluster in unstressed syllables is less than half of the duration of the same cluster in the corresponding stressed syllables, and these syllables show a pitch contour unlike that of the citation tones. Moreover, in the unstressed syllables, the last part of the sonorant cluster tends to be deleted. That is, if the underlying cluster is a diphthong, the unstressed syllable tends to show a pure vowel; if it is a VN cluster, the unstressed syllable shows a nasalized vowel, with no pure nasal band. An underlying long vowel, of course, shows only a short pure vowel in the unstressed syllables. A rule which will express the preceding could be:

(4). $[+\text{son}] \rightarrow \emptyset \ / \begin{bmatrix} \text{V} \\ -\text{stress} \end{bmatrix}$ ____

Rule (4), however, would only describe the situation adequately if it were the case that one had previously decided to specify long vowels as geminate cluster. If one had not, then another rule,

(5). $V \rightarrow [\text{-long}] \; / \; \overline{}_{\text{-stress}}$

would have to be added. Obviously, one can always formulate rules to handle the situation in which the vowels are specified as [±long] , but the question must be raised as to whether these rules are motivated, for the consequences of the specification are that one must always devise a separate set of rules to handle the long vowels separately from the sonorant clusters. Since there does not seem to be any justification for the extra rules, we shall therefore assume that the long vowels in Mandarin are abstractly represented as geminate clusters.

2.1.4 Let us now return to the problem of the features used to specify tone. Are the features to be used like those which have been proposed by Wang, features which must be assigned to but one segment in a syllable? From the preceding discussion, this kind of designation seems to ignore the basic relationship between dynamic tones and sonorant clusters.

What are the consequences of applying a system of features such as Wang's? Using these features, one could capture the relationship between the dynamic tones and sonorant clusters by saying, for example, that if a syllable were marked with the feature [+rise] , then that feature is not just a feature of one segment, but rather of the whole sonorant cluster of the syllable. However, this interpretation necessitates defining a totally new kind of feature, namely, a cluster feature. Until other evidence is produced to demonstrate that the notion of cluster feature is motivated, one would prefer to avoid such an interpretation. Alternatively, one could say that these features are distributed to all of the sonorants in a syllable, so that a sequence, say VV, would be marked $\begin{bmatrix} V \\ +\text{rise} \end{bmatrix}\begin{bmatrix} V \\ +\text{rise} \end{bmatrix}$. But not only does this appear to be extremely redundant as well as highly unmotivated, it also requires a complex phonetic interpretation of the feature [rise] , whereby within syllable boundaries, a sequence [+rise] [+rise] is not to be interpreted as two successive rises, but rather as one long rise. Outside of syllable boundaries, it must be given the interpretation of two distinct rises. Notice that a further consequence of the use of this feature in this way is the need to be able to precisely delineate syllable boundaries, the difficulties of which task are well known.

Before we continue, we will present other objections to this kind of system of features for tone. In Wang's particular system, there are not only the features [high] , [mid] , and [central] , which serve to determine five distinct pitch heights, or levels, but there are also the features [contour] , [convex] , [rise] , and [fall] . The first of the latter set, [contour] , serves only to differentiate between the static and the dynamic tones. However, it is clearly not independent of the features [rise] and [fall] , for no tone can be specified as $\begin{bmatrix} \text{+contour} \\ \text{-rise} \\ \text{-fall} \end{bmatrix}$. Any tone marked $\begin{bmatrix} \text{+contour} \\ \text{+high} \\ \text{-mid} \\ \text{-central} \end{bmatrix}$ must also be marked [+rise] or [+fall] or both $\begin{bmatrix} \text{+rise} \\ \text{+fall} \end{bmatrix}$. Nevertheless, given the specification [+contour] , one cannot predict in which direction the tone is dynamic, that is, whether it rises or falls or does both. However, one does know that the tone moves in the pitch range defined by, say, the limits high and low. That is, it moves in a range whose upper limit is high pitch and whose lower limit is low pitch. The direction of the motion is then further defined by the feature specification [+rise] or [+fall] . In the first case, this would mean that the tone moves from a point which is not high to a point that is high. However, the interpretation of the features [rise] and [fall] are then somewhat peculiar. One already knows from the specification [+contour] , that the tone is dynamic. The feature [+rise] would seem to indicate the direction of the motion. In actuality, however, it can be seen that the real function of the feature specification [+rise] is to order the realization of the pitch height features. That is, although the tone is marked [+high] , the high pitch must not be realized until the end of the production of the tone. If the tone were specified as [+fall] , then the phonetic instructions must be that the high pitch is realized immediately at the beginning of the articulation of the tone and a non-high pitch, say, low, at the end. The feature matrix $\begin{bmatrix} \text{+high} \\ \text{-mid} \\ \text{-central} \\ \text{+rise} \end{bmatrix}$, therefore, does not represent a set of physiological instructions which are to be realized simultaneously, but sequentially, as are the instructions for series of matrices within boundaries. Thus, the phonetic realization of this matrix is exactly the same as the sequential realization of the two matrices, [+high tone] [-high tone] . Is this implicit ordering device to be considered a legitimate type of distinctive feature? Before we discuss the answer to this question, let us consider one more aspect of Wang's features. The feature values [+rise] and [+fall] are not mutually exclusive in this system, for there are tones such as the isolation form of the Mandarin "third" tone, graphically represented as (∨\ ), which are supposed to both rise and fall. Thus, there can exist a tone whose feature specifications are, for example

$$\begin{bmatrix} \text{+high} \\ \text{-mid} \\ \text{-central} \\ \text{+rise} \\ \text{+fall} \end{bmatrix}$$

. In the preceding example, we saw that if the value for [fall] , was - , for example, this matrix would have been interpreted as a rising tone in the range from [-high] to [+high] , where the feature [+rise] determined the order in which the pitch heights were phonetically realized. In this example, however, we see that there are two ordering features, [+rise] and [+fall] . Which is to be realized first? Does the tone first rise from [-high] to [+high] , and then fall from [+high] back to [-high] , or does it fall first and then rise? Using only the features specified in the matrix given, there is no way to determine which is the answer to this question. Thus, Wang introduces an additional feature [convex] . A [-convex] tone first falls and then rises, and a [+convex] tone rises and then falls. Obviously, this feature serves no purpose other than to order the ordering features [rise] and [fall] . Notice also that the specification of [convex] as + or - has no interpretation except when a tone is marked as both [+rise] and [+fall]. Thus we see that this system of features contains not just features determining the order in which the pitch heights are realized, but also a feature determining the order in which the ordering features are realized.

We must now ask whether such ordering features are desirable. Do we want to say that some features of a given matrix specify the sequential order in which other features of that same matrix are realized? One consequence of this would be, of course, that unlike the distinctive features which have been established for segments, there is no immediate articulatory correlate to these features. That is, these features neither specify source, location, nor manner of articulation. Rather, they specify the order in which the actions defined by the other features in a specific matrix will be performed.[5] If one were to permit such features, one could also specify diphthongs such as /ei/, /ie/, /ou/, and /uo/, or even the triphthongs such as /iei/ or /uou/ as one segment, using features such as [rise], [fall], and [convex]. That is, using the features proposed by Chomsky and Halle in <u>SPE</u> and these three features, /ei/ could be specified as

$$\begin{bmatrix} \text{+High} \\ \text{-back} \\ \text{-round} \\ \text{-low} \\ \text{+rise} \end{bmatrix}.$$

(Unfortunately, Wang's feature [high] coincides with the feature used in SPE for tongue height, so to avoid confusion, we have taken the liberty of differentiating between the two features in this discussion by spelling the tongue height feature with a capital letter.) [+rise] in the last example would mean a transition from $\begin{bmatrix}\text{-High}\\ \text{-low}\end{bmatrix}$ tongue height to [+High] tongue height. Notice that this is simply the transition from a segment with the feature matrix $\begin{bmatrix}\text{-High}\\ \text{-low}\\ \text{-back}\\ \text{-round}\end{bmatrix}$ to the segment with the feature matrix $\begin{bmatrix}\text{+High}\\ \text{-low}\\ \text{-back}\\ \text{-round}\end{bmatrix}$. In other words, the feature [+rise] would spell out the articulatory motion required in making the transition from /e/ to /i/. Using the feature [fall], one can specify the diphthong /ie/. Using both the features [rise] and [fall] and the feature [convex], one can specify the triphthong /iei/. Obviously, one can mechanically contrive ordering features such as these to define the order of realization of articulatory motions and thus avoid having to specify a feature matrix for more than one segment. However, particularly for the vowels, this spurious saving is accomplished at the expense of ignoring that the articulatory movement of this one segment with the feature, say [+rise] , is precisely the motion which would result from articulating, for example, the vowel /e/ first and following it immediately by /i/.[6] Moreover, such a treatment would obscure the fact that this single segment behaves like a cluster, and has the same co-occurrence restrictions as other clusters, for example,/au/, which, interestingly, cannot be collapsed with just the use of the feature [rise] , rather than conforming to the constraints placed on single vowels such as /i/.

Hence we see that not only are there problems in the phonetic interpretation of the features [rise] and [fall] inside and outside of syllable boundaries, but these features are also different from other distinctive features in that they have no fixed phonetic interpretation, but serve only as ordering functions. The difficulties clearly do not just lie in Wang's particular system, which we have used only as an example, but will exist in any system of tone features which insists on using the features [rise] and [fall] . The minor objections which one can raise to the use of the feature [contour] are irrelevant. The theoretical objections must be raised with respect to introducing ordering features of the type [rise] , [fall] , and [convex] .

2.1.5 All of these complications can be avoided if a different system of distinctive features for tone is adopted. If one were to utilize the concept that a tone is specified by a sequence of pitch heights, each of which is a feature of exactly one segment, then the relationship between sonorant clusters and dynamic tones would be established. A rising tone might, for example, be defined as a sequence [-high tone] [+high tone] , where the phonetic interpretation of these pitch specifications is analogous to that of interpreting a VV sequence which is, say, /au/. In other words, the dynamic tones can be viewed as the "diphthongs" of the tone sets, with the contours being the natural result of the transition from one level to another. As with real vowel diphthongs, the transition is not a well defined point, but an interval. Moreover, such a "diphthong" concept of the dynamic tones is supported by the acoustic findings given in Kratochvíl's paper. There he states that none of the contours of the tones are strictly rising or strictly falling, but rather are what he terms "S-curves". The rising S-curve of the rising tones is a "rising contour with the curve level or with a slight 'dip' at the beginning and level or with a slight 'hump' at the end of the vowel...", and the falling S-curve for the falling tone is a "falling contour with the curve level or with a slight 'hump' at the beginning and level or with a slight 'dip' at the end of the vowel."[7]

Are there then tone triphthongs as in segmental phonology? The answer is yes, and probably these are quite as rare as three vowel sequences are. The support for a three pitch specification analysis, and hence for a representation of the syllabic nuclei of these syllables with three sonorants, comes from consideration of the data presented for the "falling-rising" tone in Mandarin in Table 2.1. One notices that all of the syllables characterized by this tone have a duration in isolation which is approximately 3/2 that of the other tones. Even more significant, however, is the fact that the vowels in these syllables are long even before a nasal. A representation of these syllables which contains three sonorants, each of which has a feature of pitch height, would then seem to be justified whenever the syllable manifests the full contour.

It therefore appears that the total duration of the sonorant cluster, beginning with the main vowel, is determined by the number of sonorants in the syllabic nucleus. Thus, the unstressed syllables are short, since these syllables have only one sonorant (recall rule (4)), the stressed syllables are approximately twice as long, and the isolation form of the "third tone" of Mandarin with three sonorants is the longest. This view gains some support from the measurements of the first pair of bisyllabic words in Table 2.2, where the syllable under examination is /lii/ (phonetically, /li/). The b/a ratio was 3.6 for this set, by far the longest ratio. However, the syllable /lii/ is also the only one in these pairs which, under stressed conditions, is pronounced with the extra long third tone. Does this imply that the lexical representation of this syllable must be /liii/? If this were the case, then one would have to say that the canonical form of the Mandarin syllable is

$$(6).\quad C(G)V(V)\begin{Bmatrix} V \\ N \end{Bmatrix}$$

rather than (2), as stated before. However, one must notice that in (6), the longest form of the syllable, $C(G)VV\begin{Bmatrix} V \\ N \end{Bmatrix}$, can only exist if the syllable is also one specified with an extra long citation tone. That is, this syllable form is uniquely associated with the extra long third tone, whereas the shorter form, $C(G)V\begin{Bmatrix} V \\ N \end{Bmatrix}$, occurs with any of the three remaining citation pitch contours.

What we wish to determine now is whether the canonical form of the Mandarin syllable must be (6), with the tone restriction on the most expanded form, and therefore, whether syllables such as /lii/ are specified in the lexicon with three sonorants in the syllabic nucleus, and with the extra long pitch contour. The support for the view that the extra long form is necessarily the lexical form arises, so far as we can tell, only from the fact that this is the phonetic shape of these syllables when uttered in isolation. One must then ask if it is the case that the lexical form of a word must correspond strictly to the form of that word when pronounced in isolation. Certainly this has never been contended for any language which is inflected, for there the lexical items can be stems which do not occur as complete "words" in speech. How does this have any bearing on Mandarin tones? It is relevant simply because the extra-long tone of Mandarin is that one which has the most distinctive sandhi forms. None of these syllables manifest the extra-long contour except when they are pronounced as isolated syllabes or when they occur at the end of a phonological phrase and are stressed. In all other positions, the pitch contour of these syllables can be described as a low level tone, except if the syllables occur contiguously in a phonological phrase. In this case, the first n-1 of n such contiguous stressed syllables assume a rising tone.[8] That is, one does not find a sequence of stressed syllables with the phonetic pitch contours as

(7). Syl low Syl low Syl low Syl rising] phon. phrase

rather, one finds the phonetic pitch contours as

(8). Syl rising Syl rising Syl low Syl rising] phon. phrase

In experiments conducted at Ohio State, it was found that the rising tone sandhi form of these syllables is not readily distinguishable from the syllables with lexically marked rising tones.[9] We therefore decided to determine if indeed the prefinal form of these third tone syllables was in any way distinctive in its duration; in particular, we wanted to know if there was any justification for stating that the structure of these syllables in non-phrase final position is C(G)VV $\begin{Bmatrix} V \\ N \end{Bmatrix}$. Thus, we recorded a series of three syllable phrases which are characterized by the stress pattern, "secondary stress, unstressed, primary stress". The position under observation was the first syllable in each phrase; thus, phrases were chosen where the tone of the first syllable ranged over three of the four basic tones. Phrases beginning with syllables in the falling tone were not used because of the already noted shorter duration of falling tone syllables in Mandarin. The nature of the third syllable in each phrase is not restricted in any way. We made Voice-prints from the recordings, and measured the duration of the first syllable and of the entire phrase. We then computed the ratio of the two durations.

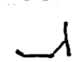
TABLE 2.3: THREE SYLLABLE PHRASES: MEASUREMENTS OF DURATION OF FIRST SYLLABLE IN S_2U S_1 PHRASES

	a) First syllable duration	b) Total phrase duration	a/b
I. 1st syllable in low level tone (sandhi form of falling-rising tone):			
sheang de heen (想得很)	42	112	.37
tzoou bu kai (走不開)	23	89	.27
nii bu doong (你不懂)	21	97	.22
leang ge ren (兩個人)	40	97	.41
II. 1st syllable in rising tone:			
yi ge ren (一個人)	23	80	.29
nan de duo (難的多)	33	86	.39
shyue bu huey (學不會)	36	105	.34
sheir de bii (誰的筆)	42	115	.37

TABLE 2.3

	a) First syllable duration	b) Total phrase duration	a/b
III. 1st syllable in high level tone:			
san ge ren (三個人)	27	76	.35
ting bu doong (聽不懂)	29	91	.32
shuo de lai (說得來)	40	99	.41

From the material presented in Table 2.3, one can see that, in the three syllable phrases where the first syllable occurs with the low toned sandhi form of the extra long citation syllables, these syllables are not distinguishable by length from the initial syllables which occur with any of the other tones. The a/b ratio, in other words, is not greater for set I, in Table 2.3, the phrases with an initial low toned syllable, than for sets II and III.

What is the lexical specification of the pitch contour of the "third tone" syllables? Let us first describe the phonetic pitch contour of these syllables when uttered in isolation. Although the contour is traditionally described as "falling-rising", it is not accurate to say that the contour is one which could be graphically represented as  . The frequency curve does not fall throughout half of the sonorant duration, nor is the fall as sharp as the rise.[10] If one looks at the spectrographic evidence, one will see that the actual contour shows a very slight drop at the beginning of the sonorant articulation, followed by a level low fundamental frequency curve, and ending with a rise to mid pitch. Graphically, this may be represented as  . We do not feel that the initial drop is distinctive, we feel that it is the natural consequence of the articulatory mechanism involved in producing the following low level tone. We will speak more about this type of phenomenon in Chapter IV. For the moment, we wish only to point out that the initial drop in these syllables is similar to the "dip" observed by Kratochvíl in the initial part of the frequency curves for syllables with the rising tone, as discussed in section 2.1.5. These "dips" have never been thought to be distinctive, likewise, we feel, the drop in the extra long syllables is not distinctive.[11]

With these phonetic facts in mind, let us now return to the problem of the lexical representation of the third tone syllables. If one were to consider the extra-long form of these syllables as the lexical form, then the phonology would have to contain rules which would delete some sonorant segment in these syllables, as well as rules which would adjust the pitch specification of the shortened syllables. These rules would have to apply every time the extra long syllables do not occur in phrase final position, for, as we have already discussed, this particular pitch configuration and the extra length are <u>only</u> found if the syllable is phrase final (and hence when uttered in isolation). Suppose that /daang/擋 <u>to obstruct</u>, which is articulated with the extra long third tone in isolation, is specified in the lexicon as / $\begin{array}{c} d \\ \\ \\ \end{array}$ $\begin{array}{c} a \\ +lowT \\ \\ \end{array}$ $\begin{array}{c} a \\ +lowT \\ \\ \end{array}$ $\begin{array}{c} ng \\ -lowT \\ -highT \end{array}$ /. We will assume momentarily that high tone and low tone are the pitch features, and that the matrix $\begin{bmatrix} \text{-high tone} \\ \text{-low tone} \end{bmatrix}$ is interpreted as a mid pitch. Whenever this syllable appears in non-phrase final position, a rule must delete one of the sonorants so that the syllable will not be phonetically realized in its extra long shape. (Recall that length seems to be directly proportionate to the number of sonorants in a syllable in Mandarin.) The sonorant to be deleted cannot be the final one, for in syllables with a nasal final, the nasal is always present. Therefore, one of the two vowels must be deleted. Let us assume that it is the second vowel. Applying the rule to /daang/, we would then derive

/$\begin{array}{c} d \\ \\ \\ \end{array}$ $\begin{array}{c} a \\ +lowT \\ \\ \end{array}$ $\begin{array}{c} ng \\ -highT \\ -lowT \end{array}$ /. Now rules to adjust the pitch contour of the derived form must apply, for when a third tone syllable occurs before another third tone syllable, the pitch contour of the first syllable is a rising one. When it occurs before any other syllable, the pitch contour is low level. Graphically, these changes may be summarized as:

(9). Third tone becomes /| before another third tone, and

(10). Third tone becomes ˩ before all other tones.

The rules specifying the changes in pitch must apply whenever the syllable is not phrase final, irregardless of whether the following syllable is stressed or not. However, if the syllable following is unstressed, recall that then the pitch of the unstressed syllable is determined from the pitch contour of the first stressed syllable preceding. In particular, if an unstressed syllable follows a syllable which has a low level tone resulting from (10), then the pitch of the [-stress] syllable is mid, or $\begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix}$. The rules which we have been discussing can be formally expressed as the following:

(11). Vowel Deletion

$$\begin{bmatrix}\text{V}\\ \text{+lowT}\end{bmatrix} \rightarrow \emptyset \;/\; \underline{\qquad} \begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix} (\#\#)\ [\text{+seg}]$$

(12). Tone Adjustment

$$[\text{+son}] \rightarrow \begin{Bmatrix} [\text{+lowT}] \;/\; [\text{+lowT}] \underline{\qquad} (\#\#)(\text{C}) \begin{Bmatrix}[\text{+highT}]\\ \begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix}\end{Bmatrix} \\ \\ [\text{+highT}] \;/\; [\text{+lowT}] \underline{\qquad} (\#\#)\ [\text{+lowT}] \end{Bmatrix}$$

(13). Unstressed syllable tone assignment

$$\begin{bmatrix}\text{V}\\ \text{-stress}\end{bmatrix} \rightarrow \begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix} \;/\; [\text{+lowT}]\,[\text{+lowT}]\,(\text{C}) \underline{\qquad}$$

One must specify [+seg] in rule (11) in order to determine that the syllable is not phrase-final, that situation presumably being marked by some kind of boundary. In rules (11) and (12), word boundaries may occur in the environment, for the syllables conditioning the change need not be clitics. However, in (13), no word boundaries can occur, for this rule only applies to unstressed syllables which are clitics phonologically.

Rules (11)-(13) will account for the facts, but they are cumbersome and do not capture all of the possible generalizations. Consider the pitch contour of the verb /daang/, cited above, when it is uttered in isolation, and the pitch contour of the verb phrase /daang + le/, where /le/ is the completive aspect marker. When /daang/ occurs isolated, the pitch contour is $[+\text{lowT}]\,[+\text{lowT}]\begin{bmatrix}-\text{highT}\\-\text{lowT}\end{bmatrix}$. In the phrase, however, since it occurs before another syllable, it is shortened by (11) and the pitch configuration is adjusted by the first part of (12). The particle /le/ is a clitic, and therefore subject to the pitch assignment rule, (13). The pitch contour for the word /daang le/ is $[+\text{lowT}]\,[+\text{lowT}]\begin{bmatrix}-\text{highT}\\-\text{lowT}\end{bmatrix}$, identical to that of the verb in isolation. The identity of these two contours has been remarked on in the literature, and observed in laboratory experimentation by Wang and Li.[12]

This particular contour is restricted in its occurrence to words containing third tone syllables fulfilling certain conditions; namely, either the word consists only of the syllable, and is phrase final, or the word consists of the low level tone variant of the third tone syllables before an unstressed syllable. Is it not possible for the phonology to capture these constraints in a more explicit manner? If we do not consider the extra long form of the third tone syllables as the lexical form, then we can both capture the relationship and present phonological rules to give the proper forms which are less complicated than (11)-(13). Suppose one were to consider the lexical form of /daang/ as $/d\ \underset{+lowT}{a}\ \underset{+lowT}{ng}\ /$. Then the following rules will account for all forms:

(14). $[+lowT] \rightarrow [+highT]\ /\ [+lowT]\ \_\_\ (\#\#)\ [+lowT]\ [+lowT]$

(15). $\emptyset \rightarrow \begin{bmatrix} V \\ +lowT \\ \alpha F_i \end{bmatrix} / \left[\begin{bmatrix} V \\ +lowT \\ \alpha F_i \end{bmatrix} \_\_ \begin{bmatrix} +son \\ +lowT \end{bmatrix} \right]_{phon.\ phrase}$

where $\{F_i\}$ is the set of segmental features

(16). $[+son] \rightarrow \begin{bmatrix} -highT \\ -lowT \end{bmatrix}\ /\ [+lowT]\ [+lowT]\ \_\_$

Rule (14) gives the special rising tone form for a series of low tone syllables,[13] rule (15) gives the extra long phrase final form, and rule (16) adjusts the tone of both the final sonorant in the extra long phrase final forms, as well as of the sonorant of the particles. This last set of rules is not only simpler, but it also captures the notion that the extra long forms are a special form, and that the configuration $[+\text{lowT}]\,[+\text{lowT}]\begin{bmatrix}-\text{highT}\\-\text{lowT}\end{bmatrix}$ arises only when the syllables with lexical low level tones are in the environment of a third sonorant enclosed by some tight phrase boundary.

If one accepts this analysis of the third tone in Mandarin, then it is possible to state that the syllable structure is simply $C(G)V\begin{Bmatrix}V\\N\end{Bmatrix}$, as given in (2), where G is a prevocalic glide, and is therefore [-syllabic]. The Mandarin tones can thus be described as [+highT] [+highT], the high level tone (˥), $\begin{bmatrix}-\text{highT}\\-\text{lowT}\end{bmatrix}$ [+highT], the rising tone (˧˥), [+highT] [+lowT], the falling tone (˥˩), and [+lowT] [+lowT], the traditional third tone.

<u>2.2.0</u> We have now seen that at least for one of the Oriental languages, a segmental analysis of tone is valid. It would then appear that, unless the structure of Mandarin is shown to be radically different from all of the other languages of the Orient, the segmental

analysis should hold for those other languages. We have done some spectrographic examination of Cantonese and find that the segmental analysis certainly holds there. Furthermore, a look at Abramson's findings in his study, The Vowels and Tones of Standard Thai, would indicate the same for Standard Thai. Thus, we contend that all those languages whose tones have traditionally been described in terms of the features [rise] and [fall], may also be tonally described as having dynamic tones which are sequences of pitch heights, each of which is a feature of some sonorant segment. If the tonal phenomena of all languages are essentially the same, and therefore should be describable using only one set of distinctive features, our choice then seems to lie between a system which uses features designating contour tones as features of one segment, and a system which describes contour tones as sequences of pitch heights. We must now ask if there are any objections to a system using the features [rise] and [fall] other than the theoretical objections stated earlier. Specifically, we must ask whether it will be the case that all those languages whose dynamic tones have previously been described as a sequence of two pitch heights can be adequately described using these other features.

In order to determine this, we shall now examine the tonal phenomena of Northern Tepehuan in detail. It will be shown that tone is not lexical in this language, but rather is entirely predictable. In our discussion, we will describe the dynamic tones as pitch sequences and formulate all rules accordingly. We will then attempt to reformulate them using the features [rise] and [fall] and discuss the repercussions of this reformulation. Notice that in order for our statement that the tones of all languages are to be treated as the same phenomenon to hold, it is necessary that the same set of features be capable of adequately describing the tones of Northern Tepehuan, a non-lexical tone language, as well as those of Mandarin.[14]

2.2.1 Northern Tepehuan, a Piman language spoken in southwestern Chihuahua, Mexico, has been described by Dr. Burton Bascom as a tone language with two tones, high and low, hereafter abbreviated H and L, and six tonal sequences.[15] The tones are also said to be phonemic, that is, marked in the lexicon. Four of the tone sequences are restricted to syllables with long vowels or diphthongs; HH, HL, LH, and LL, or in other terminology, long high level, falling, rising, and long low level. The other two, which are the unit sequences H and L, occur on syllables with short vowels. The vocabulary of this language, however, is not monosyllabic. The tones are further described as being subject to sandhi conditioned by morphological suffixing.[16]

A cursory glance at the vocabulary of Northern Tepehuan would seem to reveal a random assignment of tone configurations to lexical items, thus supporting the thesis that tone is phonemic. Examples of forms in which these tonal contours occur are the following: /bánai/ <u>coyote</u>, /daáka/ <u>nose</u>, /kookóso/ <u>they sleep</u>, /móo/ <u>head</u>, /víótai/ <u>to vomit</u>, /góógoši/ <u>dogs</u>, /dagívĩñai/ <u>he kneads</u>, /kukúntagi/ <u>they take a husband</u>, /kóva/ <u>forehead</u>, /maákai/ <u>he gives</u>.[17]

If one examines the lexicon more carefully, however, it will be seen that the system of tone assignment is not totally random. If one considers the rising and falling tones not as prosodic units, but rather as sequences of pitch levels where rise = LH, and fall = HL, then the following statement can be made about the restrictions on the occurrence of the high tone. For all forms, high tones never occur on any but the first or second vowel clusters irrespective of the length of the form. The terms "first" and "second" refer to the clusters as they are counted beginning with the left-most boundary of the stem, not of the form. Prefixes play no role in tone assignment. Hence, in a phrase like /iñkoovába/ on my forehead, /iñ-/ is the possessive pronoun my, which behaves as a prefix on the noun. Thus, although the high tone is found on the third vowel cluster of the form, it is still only on the second cluster from the boundary of the stem.

2.2.2 These are not the only regularities, however. If one looks at the nouns, it becomes evident that for most of them, the tone shape of the singular is not identical with the tone shape of the plural. Examples of various noun forms are the following: /báhi/ tail, /baabáhi/ tails, /bahíd^{y}ɨ/ his tail, /bánai/ coyote, /baabánai/ coyotes, /daáka/nose, /daadáka/ noses, /daakádɨ/ his nose, /nakásɨrai/ scorpion, /nanákasɨrai/ scorpions, /tukúrai/ owl, /tutúkurai/ owls. The method for forming plurals is by reduplicating the first syllable. If one considers this reduplication to be a process of infixing rather than prefixing, then the infixed copy of the first syllable can be said to always appear with a short vowel, irrespective of the length of the vowel originally.[18] Thus, one has both /báhi/ and /baabáhi/, /daáka/ and /daadáka/, where the second syllable in the plural forms is considered the reduplicated syllable. If the singular form is bisyllabic, then the vowel of the first syllable is lengthened if it is not already long. Hence,

/bánai/ - /baabánai/ and /daáka/ - /daadáka/. For nouns which in the singular are trisyllabic or longer, the plural is formed as for the shorter nouns, except that the vowel of the first syllable is left unchanged. Hence, /nakásɨrai/ - /nanákasɨrai/ and /tukúrai/ - /tutúkurai/. Looking at just the nouns which are bisyllabic in their singular forms, one might think that there is a rule of accent shift which moves the high tone from the first syllable to the second when the form is made trisyllabic or longer (through plural formation or suffixation). Hence, /daáka/ but /daadáka/ and /daakádɨ/, as well as /kií/ *house*, /kiíki/ *houses*, but /kiikíši/ *are they houses?* Thus, if one wished to hold the view that the tones form a part of the phonological specification of each lexical entry, and presuming that only the singular form of any noun occurs in the lexicon, one might be tempted to say that there is a rule of high tone shift for nouns which moves a high tone to a vowel in the next syllable if the form is made trisyllabic or longer through any morphological process. However, that such a rule of tone shift cannot account for all nouns is obvious from consideration of the last two sets of examples, namely, /nakásɨrai/ beside /nanákasɨrai/ and /tukúrai/ beside /tutúkurai/. Here the singular is itself polysyllabic, and the high tone is found on the second syllable in both the singular and plural forms. It is in fact the case that for all noun forms, the high tone occurs on the second syllable if the form is trisyllabic or longer, and on the first syllable otherwise, irrespective of whether the form is singular, plural, or suffixed. There are two sets of noun forms which would appear to be exceptions to this rule, some examples of which have already been given, e.g. /góógoši/, but these will be discussed later.

Since plural formation and suffixation such as with the third person possessive pronoun /dɨ/ lengthens a form, the location of the high tone in the simple singular form may differ from that of the plural or a possessed form. Thus, it is seen that for the nouns, if tone is considered lexical, and if only the singular form appears in the lexicon, then a complex set of rules must be established to relate, among other things, the tonal contour of plurals to singulars.

2.2.3 As far as verbs are concerned, there are two morphological processes which are relevant to the discussion of tone. The first is the formation of the usitative or habitual stem. This stem is formed by reduplication of the first syllable. Thus, from /kiísa/ to scratch, one gets /kikísai/ he keeps on *scratching*.[19]

The second process is the formation of the perfective stem, which is accomplished by truncating the final syllable of the base form. If the base form is bisyllabic, then there is a further change in the vowel of the initial syllable. If the consonant which begins the truncated syllable is a [-coronal] consonant, then the vowel of the initial syllable is lengthened, if it is not already long. If the consonant is [+coronal], and the preceding vowel is already long, then the long vowel becomes a diphthong with an /i/ glide. Otherwise, the glide /i/ is added. Thus, one has /maákai/ *he gives*, /maá/ *he gave*, /taánɨi/ *he asks for*, but /taí/ *he asked for*. For base forms which are trisyllabic or longer, the formation of the perfective stem causes no lengthening of the vowel in the preceding syllable. Thus: /soobídʸai/ *to head off* and /soóbi/ *he headed off*. It is because of the similarity of the behavior of the long vowels with the real diphthongs in these instances, among others, that has motivated the representation of the long vowels as clusters of identical vowels. By utilizing such a representation, the tones on the long vowels as well as on the diphthongs, can be represented as sequences of the tones which occur on short vowels, thus reducing the number of features necessary to distinguish the various tones to one, [high tone] , where [+high tone] is high, and [-high tone] is low.

Other examples of verb forms are: /kunátai/ *to take a husband*, /kukúntagi/ *they take a husband*, /muúli/ *to break stick*, /mɨ́rai/ *he runs*, /mɨɨ́/ *he ran*, /koóso/ *he sleeps*, /kookóso/ *they sleep*, /gikúdai/ *he whistles*, /dagíviñai/ *he kneads*. One sees immediately from these and the previous examples, that the same relations hold for the verb forms as for the nouns; the high tone occurs on the second syllable if the form is trisyllabic or longer, and on the first syllable otherwise.

2.2.4 While with the nouns it was clear that if the lexicon specified the tones of the singular forms, then a complex set of rules would have to be contrived to account for the behavior of the singular forms which are bisyllabic, as opposed to those which are trisyllabic or longer, if the forms were changed by, say, possessive suffixing or plural formation, with the verbs, one sees that the situation is even worse. Not only would base forms have the same problems as the nouns regarding plural as opposed to singular forms, but now the tones of the different morphological stems, other than the plural, must also be related to one another by a set of rules.

On the other hand, since tone appears to be so directly connected to length of form for both nouns and verbs, if one abandons the notion that tone must be lexical, then the relation of high tone position to length of form may be expressed for all morphological forms by a simple set of rules. Such a set of rules might be the following:

(17). $V \rightarrow +H \; / \; _{\text{Stem}}\big[\, C_o(V)\_\_\_C_o(V)\#\#$

(18). $V \rightarrow +H \; / \; _{\text{Stem}}\big[\, C_o(V)VC\_\_\_CV$

We assume here that the feature [high tone], abbreviated as [H] in all of the Northern Tepehuan rules, is a normal part of the feature matrix for each sonorant. Thus, in Northern Tepehuan, each lexical entry has the unmarked value of this feature, - , for every sonorant segment. We therefore do not need a rule assigning [-high tone].

Rule (17) takes care of the monosyllabic and bisyllabic forms. The environment, $C_o(V)\_\_\_C_o(V)\#\#$, indicates that if the vowel is long, or a diphthong, the tone is placed on the second of the two vowels, thus giving a rising contour. In this rule, ## indicates that the form to which it applies cannot be more than two syllables long. Rule (18) places the high tone on the second syllable of longer forms. Since anything which might be prefixed does not count as the first or second syllable to which the high tone might be assigned, the left bracketing indicating <u>Stem</u> is necessary. Clearly, the rules must be ordered as given.

While these two rules actually account for a great deal of the data, one must recall that the tone system on long vowels is described as consisting of four sequences. The above rules only account for two, LL and LH. (The symbol L stands for the feature specification [-high tone] .) What about HL and HH? The first of these two configurations, HL, is very restricted in its occurrences. It can only appear on an ostensibly monosyllabic form, and it is the case that if the form becomes lengthened through suffixation, the contour is changed to HH. Thus, one finds /móo/ <u>head</u>, but /móódɨ/ <u>his head</u>; /ái/ (as in /kašiš ái/ <u>did he already catch up to him?</u>), but /áíyata/ <u>he caught up to him</u>. How can these items be accounted for? They are, to be sure, relatively rare, but they are sufficiently numerous to cause one to seek some property by which to characterize the set if one wishes to consider them as exceptions. Before attempting to state that characterization, however, a discussion of the other instances of the HH configuration on the first vowel cluster is warranted. As just shown, some of the forms which show a HH contour are also found with a HL contour, if the form can be reduced to an ostensibly monosyllabic one. However, there are also cases such as /báágai/ eagle, /vúápai/ he brings, /bíúgimu/ to be hungry, and /ááduñi/ relatives, which are not analyzeable as an ostensibly monosyllabic base with a suffix as, say, /áíyata/, which is the stem /ái/ and the suffix /-yata/. From an inspection of these forms, one cannot see any similarity which would allow them to be classed as a single set of exceptions. If one looks at the historical evidence, however, then one striking characteristic immediately comes forth. The vowel clusters in forms which show either a HH or HL tone contour and which are the first cluster in the stem, can all be traced back to situations in which the members of the cluster were vowels of distinct syllables. That is, for Proto-Tepiman, Bascom has reconstructed the forms /*baʔaga/ for /báágai/, /*moʔo/ for /móo/, /*ʔuʔapa/ for /vúápai/, /*bihugimu/ for /bíúgimu/ and /*hahaduni/ for /ááduñi/. The instances of /ʔ/ and /h/ which Bascom reconstructs for Proto-Tepiman, which are realized as zero intervocalically in Northern Tepehuan, arise from both Piman and Uto-Aztecan laryngeals.[20] The Uto-Aztecan /*ʔ/ became the Piman /*ʔ/, which, in Northern Tepehuan, is zero. The Uto-Aztecan /*h/ became zero in Piman.

The Piman /*h/ which stems from Uto-Aztecan /*s/, is the Proto-Tepiman /*h/, which sometimes became zero in Northern Tepehuan. Thus, one can see that by the loss of an intervocalic laryngeal, the vowels of adjacent syllables became what superficially appear to be normal vowel clusters. However, if one were to assume that in some way, the previous existence of a non-vocalic segment between these vowels were retained in the synchronic phonology, then the placement of the high tone on the ostensibly monosyllabic forms such as /móo/ is correctly predicted by rule (17).

One might propose that this "remembrance of the lost laryngeal" is effected by positing the existence of some underlying phoneme /L/ which would occur in the lexical specification of all these items. For example, the underlying form of /móo/ would then be /moLo/. The nature of this segment might be characterized as being some laryngeal which does not otherwise occur in Northern Tepehuan. Then the tone contour of /móódɨ/ would be the result of the following rules along with rules (17) and (18).

(19). $V \rightarrow +H \ / \ \_\_\_\_ \ L\begin{bmatrix} V \\ +H \end{bmatrix}$

(20). $L \rightarrow \emptyset$

Thus, /móódɨ/ would have the form /moLodɨ/ before the application of rule (17). Rule (17) would then put a high tone on the second vowel, yielding /moLódɨ/. Rule (19) states that a vowel which occurs before the segment /L/ followed by a high toned vowel is itself assigned the high tone. Applied to /moLódɨ/, rule (19) gives /móLódɨ/. Rule (20) simply deletes the segment /L/, giving the final shape, /móódɨ/.

Is the only solution to this tone problem the one which posits the existence of a hypothetical laryngeal and, if not, is it the best one? The answer to both questions is seemingly no. Before discussing an alternative, however, some more forms must be presented. The noun /gogóši/ dog has as its plural form /góógoši/. The plural verb stem for stand is /gúúka/. Why does the plural form of dog have this shape? According to the rules for plural formation, the plural should be /gogógoši/. Likewise, the verb form should be /gugúka/. The same forms in related languages, namely, Southern Tepehuan and Upper Piman, are /goggoš/ and /gogogšɨ/, /guguuk/ and /gɨgukɨ/, respectively. For Lower Piman, the form for dogs is /googiš/. From this one can either conclude that these forms are completely irregular for both Northern Tepehuan and Lower Piman, or one might state that Northern Tepehuan, and possibly also Lower Piman, has a rule in its phonology which is not in the phonologies of either Upper Piman or Southern Tepehuan. Such a rule would simply state that if three velar consonants begin consecutive syllables in any given form, then the second of such consonants is deleted. This then permits the plural for dog to be regular and the form for the verb to stand to be irregular only in that it has a suppletive plural base; the morphological process of forming the plural once given the suppletive base would then be the normal process. Since the morphological processes of all of the Piman languages are extremely similar, the addition of a velar deletion rule preserves the relationship between the different languages as well as preserving the morphological regularity within Northern Tepehuan.

The tone contour for /góógoši/, for example, could then be accounted for by changing the form of the velar deletion rule so that instead of deleting the velar, it changes the velar to the laryngeal /L/ in exactly the same environment in which the velar would have before been deleted. By ordering this revised form of the velar deletion rule before rule (19), the proper tone contour would emerge. Rule (20) would then delete all occurrences of /L/. The only problem with this solution, therefore, is the original problem of justifying the existence of the laryngeal /L/.

Because the tonal behavior of those vowel clusters which arise from a synchronic loss of an intervocalic consonant is identical to the behavior of those which arose from an historic loss of a non-vocalic intervocalic segment, and because this behavior is different from that of genuine long vowels and diphthongs, it would seem that there should exist a method, other than that of positing the existence of a hypothetical laryngeal, which would distinguish vowel clusters which represent two syllabic nuclei as opposed to one, as do long vowels and diphthongs.

2.2.5 Such an alternative solution to this tonal problem which distinguishes between these types of vowel clusters can be contrived using the feature [syllabic].[21] This feature can be roughly defined as a feature of all segments, with the + value limited to [+sonorant] segments. By using this feature in place of [vocalic], and stating that all segments which have all the feature specifications which otherwise determine a sonorant can be either specified as [+syllabic] or [-syllabic], one then sees even more clearly the formal similarity of long vowels and diphthongs. That is, since reasons have already been offered for considering long vowels in Northern Tepehuan to be VV sequences of identical vowels, then both long vowels and diphthongs consist of a syllabic vowel followed by a non-syllabic vowel, or in other words, a glide. Thus, the similar tonal behavior of /maá/ <u>he gave</u>, /taí/ <u>he asked for</u>, /daáka/ <u>nose</u>, and /taíši/ <u>is it fire?</u>, is readily explained. In monosyllabic and bisyllabic forms, the high tone occurs on the glide of the first syllable if there is one.

On the other hand, suppose that for forms like /v́úápai/,/móo/, and /góógoši/, the vowel sequences are not sequences of a syllabic member followed by a non-syllabic member, but rather that both vowels are syllabic. The contention here is that vowel clusters which arise either by an historic loss or by a synchronic loss of a non-vocalic intervocalic segment, may retain the property of being different syllabic nuclei. In this manner, the lexical specification of the vowel cluster in /móo/ and in /v́úápai/ would state that both vowels are syllabic, and hence different, as evidenced by the tonal behavior of these forms, from diphthongs and long vowels. Likewise, when the velar is deleted from the form /gogogosi/, by means of the unrevised velar deletion rule, the vowel following the deleted velar remains a syllabic vowel. Can rules (17) - (20) be reformulated utilizing the feature [syllabic] rather than the hypothetical laryngeal /L/?

Consider the following rules:

(21). $V \rightarrow +H \;/\; {}_{Stem}\left[\, C_o(V)\left(\begin{bmatrix}V\\ -syl\end{bmatrix}\right)_o^2 C_o \;\overline{\langle +syl\rangle}_a\; \left\langle C_o \begin{bmatrix}V\\ +syl\end{bmatrix}\right\rangle_b\right.$

where $\langle\,\rangle_a$ and $\langle\,\rangle_b$ are interpreted as:

If a then b

(22). $V \rightarrow +H \;/\; \overline{+syl}\; \begin{bmatrix}V\\ +H\\ +syl\end{bmatrix}$

Rule (21), by using the angle bracket notation, places the high tone on the second syllabic vowel of a form if there is another syllable following. If the form is monosyllabic or bisyllabic, the high tone is placed on the last vowel of the first syllable. Notice that in a bisyllabic form such as /bánai/, the high tone can be placed on the syllabic vowel of the first syllable because there is another syllable following. However, the high tone can never be placed on the syllabic vowel of a monosyllabic form because of the restriction expressed by the angle brackets. The brackets are to be interpreted as "if a then b".[22] Thus, the environment using the angle brackets is an abbreviation of the following three strings:

(23). $C_o(V)\left(\begin{bmatrix}V\\ -syl\end{bmatrix}\right)_o^2 C_o \;\overline{+syl}\; C_o \begin{bmatrix}V\\ +syl\end{bmatrix}$

(24). $C_o(V)\left(\begin{bmatrix}V\\ -syl\end{bmatrix}\right)_o^2 C_o \;\rule{2em}{0.4pt}\; C_o \begin{bmatrix}V\\ +syl\end{bmatrix}$

(25). $C_o(V)(\begin{bmatrix} V \\ -syl \end{bmatrix})_o^2 \; C_o$ ———

This expansion is somewhat deceptive, for (24) and (25) should represent the environments for bisyllabic and monosyllabic forms respectively. However, the expanded forms (24) and (25) seem to apply to trisyllabic and bisyllabic forms. The deception lies in the fact that although (24) and (25) theoretically allow a consonant to exist before the vowel under consideration, this case cannot in fact exist. That is, if a consonant appears before the vowel in question, then that vowel must be [+syllabic], the situation described in (23). If the vowel being considered is [-syllabic], then no consonant may appear before it, for there are no glides which follow consonants in Northern Tepehuan. Thus, the real expansion of the schemata given in (21) is:

(26). $C_o(V)(\begin{bmatrix} V \\ -syl \end{bmatrix})_o^2 \; C_o \; \overline{\underline{\quad}}_{+syl} \quad C_o \begin{bmatrix} V \\ +syl \end{bmatrix}$

(27). $C_o(V)(\begin{bmatrix} V \\ -syl \end{bmatrix})_o^2 \quad \underline{\quad\quad} \quad C_o \begin{bmatrix} V \\ +syl \end{bmatrix}$

(28). $C_o(V)(\begin{bmatrix} V \\ -syl \end{bmatrix})_o^2 \quad \underline{\quad\quad}$

From the expanded forms (26) - (28), it is clear how rule (21) applies. Given the forms /daadáka/ and /tukúrai/, the second syllabic vowel is assigned the high tone, since these forms fulfill the conditions of the environment as expanded in (26). For forms such as /móo/, the first vowel will be assigned the high tone since both vowels are [+syl] , but only the first one satisfies the condition stated by the angle brackets, that there is another syllabic vowel following. /daáka/, on the other hand, is assigned the high tone on the last vowel of the first syllable since it fulfills the conditions in (27). All monosyllables such as /taí/ will have the high tone placed on the last vowel as seen from (28).

Rule (22) simply states that when a syllabic vowel immediately precedes another syllabic vowel which has been assigned the high tone by (21), then that vowel is also assigned the high tone. Thus, in a form such as /móódɨ/, rule (21) applies to give /moódɨ/, where the second syllabic /o/ gets the high tone because the form is trisyllabic. Rule (22) then applies to give /móódɨ/. Notice that in a form such as /vúápiava/ he does bring, where the first two vowels are [+syl] , rule (21) will not give the incorrect assignment, /vuapíava/*, because the environment of the rule specifically states that all but one of the vowels preceding the one under consideration must be [-syl] . Since the first two vowels in /vúápiava/ are syllabic, (21) will apply to give /vuápiava/. (22) will then give

/vúápiava/. Notice that (22) does not apply to forms like /daáka/, where the vowel cluster consists of a syllabic vowel followed by a [-syl] vowel or glide.

In the environment of rule (21), one will notice that the vowel under consideration may be preceded by at most one consonant. This restriction must be stated, for one finds that there are forms such as /ást^{y}añi/ <u>throw it out</u>, /síspunai/ <u>it explodes</u>, /táškali/ <u>tortilla</u>, and /sástuduakana/ <u>clever (plural)</u>. While all of these forms are trisyllabic or longer, all have the high tone on the first syllable, an apparent violation of the rule for high tone placement on polysyllabic forms. However, as is obvious, each of these forms is characterized by the existence of a consonant cluster before the second syllable. Thus, if one examines carefully the environment of rule (21), one sees that the high tone is placed on the second syllabic vowel only if that vowel is preceded by zero or one consonants. For this reason, the vowels following the consonant cluster in the preceding forms cannot be assigned the high tone. Rather, it will be the last vowel preceding the cluster which is assigned the high tone.[23] Because of the existence of forms such as /mumúrtaimi/ <u>it is tasseling</u>, and /kukúntagi/ <u>they take a husband</u>, it is clear that the existence of a consonant cluster only affects the placement of the high tone if the clusters occur before the second syllable. The question which must now be raised is whether these clusters are underlying clusters, i.e. listed in the lexicon as such, or whether they arise from the deletion of a vowel, and if so, can the rule of the phonology which deletes the vowel be ordered before rule (21). We must determine the answer to this question, for if we cannot assume that the clusters exist before the application of rule (21), then we will have to complicate the tone rules to adjust the tone assignment in the event that the vowel to which the high tone has been assigned is deleted.

There are several arguments in favor of deriving all such clusters, not the least important of which is, of course, the fact that then one does not have to state complicated constraints on where such clusters may occur as part of the morpheme structure rules. These clusters never occur initially, and, except in one well-defined instance, they cannot occur in a form at the beginning of the ultimate syllable.[24] However, since syllables which are morphologically attached suffixes may count as the final syllables of the form, it would appear that at least for some forms, a rule is needed to delete the vowel. Thus, one has the form /kunátai/ to take a husband, which consists of the noun /kúna/

husband, and the verbal suffix /-ta(i)/. There is also, however, the form /kukúntagi/ they take a husband, which consists of the plural of the stem /kunátai/ with the suffix /-gi/. Since a rule would be necessary to delete the vowel for forms such as /kukúntagi/ and /sastuduakana/ (the plural form of /satuduakana/), and since those items such as /táškali/ share exactly the same constraints as to which consonants can occur in the clusters as well as to the position of the cluster with respect to ##, it is thus proposed that all clusters be derived by rule. The rule can be ordered before (21), for there is nothing in either rules (21) and (22) or in any subsequent rules which depends on the vowel being present. Rule (21), therefore, because of its constraint on the number of consonants which can occur before the second set of vowels in a form, will correctly assign the tone in forms with consonant clusters.

The next point to be discussed is the behavior of monosyllabic forms when they are suffixed. Examples of such forms are the noun /taí/ fire, and the perfective stem /saí/ (as in /kašiš saí/ did he already get stuck?) Suffixes which occur with nouns include /-ši/, /-kɨdɨ/, /-ana/, /-aba/. Suffixes occurring with verbs include /-ši/, /-ata/, /-nɨ/, /-ava/, and /-t^{y}iki/. Thus one has the forms /taíši/ *is it fire?*, /taiyána/ *at the fire*, /taiyába/ *on the fire*, /saíši/ *did he get stuck?* /saínɨ/ *I got stuck*, /saiyáta/ *he got stuck*. From the tonal contour of forms such as /móókɨdɨ/ *with his head*, it has already been seen that the suffixes are capable of being counted in determining whether or not a form is polysyllabic. The form /móókɨdɨ/ consists of the stem /móo/ and the suffix /-kɨdɨ/. When the noun occurs by itself, the tonal contour is /móo/ as predicted by rule (21). However, when the suffix /-kɨdɨ/ is attached, the form is no longer bisyllabic, but multisyllabic, so that rules (21) and (22) apply to give /móókɨdɨ/. Looking at the various forms of /taí/ and /saí/ above, one sees that apparently the suffix itself can contain a vowel which is assigned the high tone. For example, this is the case in /taiyába/, which is analyzeable as the stem /taí/ and the suffix /-(y)aba/. However, it is obviously not the case that all suffixes contain vowels which can be assigned the high tone by (21), even if they occur with a monosyllabic stem, for one also finds the forms /taíkɨdɨ/ *with the fire*, and /saít^{y}iki/ *he did get stuck*. Here, although the forms are trisyllabic, and although only a single consonant occurs before the second syllabic vowel, the high tone does not occur on the second syllable. However, all forms such as /taíkɨdɨ/ and /saít^{y}iki/ are uniquely characterized by the fact that the consonant of the second syllable is preceded by a /+/ boundary. Thus it is the case that the high tone can occur on a vowel in a suffix attached to a monosyllabic stem only if the vowel in the suffix is not the vowel of the last syllable of the form, and the suffix does not begin with a real consonant. The first condition is exactly the same condition which holds for any type of form, namely, the high tone never appears on the final syllable unless the form is monosyllabic. The second condition, however, is unique to suffixes.

Before attempting to formulate rules to handle the special case of suffixes, a comment must be made about the suffixes. In the preceding sets of examples, the suffixes beginning with a vowel, for example, /-ata/, have been seen to occur in a variant shape with a prevocalic glide inserted, for example, /-yata/. This variant shape of the suffix is realized phonetically when the stem to which it is attached ends in the vowel /i/, either syllabic or non-syllabic, and when the stem contains no consonant other than the initial consonant. Thus, when the suffix /-ata/ is attached to the stem /saí/, the resultant form is /saiyáta/. When it is attached to the stem /ái/, the resultant form is

/áíyata/. We will discuss the nature of the glide insertion rule later, and show that, in fact, one must posit the existence of this glide in all forms where the stem to which the suffix is attached contains no consonant other than the initial consonant, even if the stem does not end in /i/. The glide insertion rule does not have to precede (21). Thus, we may assume that (21) applies to /tai + ata/, for example, and not to /tai + yata/.

Rule (21), when applied to /tai + ata/ and /tai + kɨdɨ/, will give /tai + áta/ and /tai + kɨ́dɨ/. We therefore need rules to adjust the tone assignment in /taikɨ́dɨ/ to /taíkɨdɨ/ which will not affect /taiyáta/. Such rules can be:

(29). $V \rightarrow +H \ / \ \_\_\_\_ + [+cons] \begin{bmatrix} V \\ +H \end{bmatrix}$

(30). $V \rightarrow -H \ / \ + [+cons] \ \_\_\_\_$

(29) shifts the high tone to the last vowel preceding the morpheme boundary whenever the vowel in a suffix beginning with a consonant has been assigned the high tone by (21). (30) states that the vowel following the initial consonant of a suffix is always low toned.

So far, we have not given any real argument against the hypothetical laryngeal, for all of the data presented up to now can be handled by rules using the hypothetical laryngeal /L/, rather than the feature [syllabic]. Thus, rule (21) can be reformulated as

$$(31).\quad V \rightarrow +H \;/\; {}_{\text{Stem}}\Big[\, C_o(V)(V)\ \langle VC\rangle_a \;\_\_\_\; \langle CV\rangle_b$$

where, if a then b

In (31), the angle brackets are interpreted as in (21). The environment, therefore, is an abbreviation of the following three strings:

(32). $C_o(V)(V)VC\_\_\_CV$

(33). $C_o(V)(V)\_\_\_CV$

(34). $C_o(V)(V)\_\_\_$

In order to argue against the existence of the hypothetical laryngeal, we must show that the only purpose it serves is to make syllabicity non-distinctive in the lexicon. That is, if we can show that the first tone assignment rule must mention the feature [syllabic], then it will be the case that the hypothetical laryngeal is needed for no rule, since the purpose of its existence is to avoid having to mention syllabicity in the environment of any rule. Rule (31), in conjunction with rules (19) and (20), will apply to the forms /moLo/, /vuLapai/, and /sai + ata/ correctly to give /móo/, /vúápai/, and /saiyáta/ (after the glide insertion rule). However, let us consider the form /saiyápɨta/ <u>you got stuck</u>, which consists of the perfective stem /saí/ and the second person singular suffix, /-apɨta/. Here the suffix is itself trisyllabic and the form contains four syllables. However, if (31) is applied, the result will be /sai + apɨ́ta/, an incorrect assignment. The point of this example is simply that with a form such as /saiyápɨta/, one cannot apply (31) and get the right result, because there is no way of determining that this form contains four syllables. That is, (31) determines the number of syllables in a form, and therefore the vowel on which to place the high tone, by counting consonants. It therefore treats /saiyápɨta/ as if it were a trisyllabic form. The right results are given for a form like /saiyáta/ only because this rule applies to the form as if it were bisyllabic, like /daáka/, and therefore places the high tone on the vowel immediately preceding the second consonant, which is coincidentally the right vowel. That is, the correct assignment is made in /saiyata/ completely accidentally. Because there are trisyllabic suffixes like /-apɨta/ which begin with a vowel and can occur with a monosyllabic stem, we see that a rule such as (31) cannot apply correctly. No rule which predicts syllabicity from consonants can apply correctly, for there is no consonant preceding suffixes like /-apɨta/, and there are no reasons for inserting one there.[25]

On the other hand, notice that (21) applies correctly to /saiyápɨta/, for it states that the high tone is placed on the second syllabic vowel whenever another syllabic vowel follows, <u>irrespective</u> of the nature of the preceding segments. Only a rule which mentions syllabicity explicitly can apply correctly to these forms. Since the first tone assignment rule must make mention of the feature [syllabic] , what purpose does the hypothetical laryngeal serve? The first tone rule would have to mention the feature [syllabic] even if the hypothetical laryngeal existed. It therefore follows that the only motivation for /L/ arises from the desire not to use syllabic distinctively in the lexicon. The choice then lies between a lexical specification /moLo/ and /m $\underset{+syl}{\text{o}}$ $\underset{+syl}{\text{o}}$ /. In the first case, there would have to be a rule stating that all vowels after consonants are syllabic, and another rule deleting the /L/. Both of these rules must apply before (21). We do not feel that the ability to avoid using syllabic distinctively in the lexicon is sufficient justification for the existence of /L/.[26] We therefore abandon the notion of the hypothetical laryngeal.

In rule (21), the environment was described as

$_{\text{Stem}}\, C_o(V)\left(\begin{matrix} V \\ -\text{syl} \end{matrix}\right)_o^2$ .... , where the vowel of the first syllable has been optionally followed by two glides. This has been done to permit forms such as /doáí/ (as in/kašiš doáí/ did he already get well?) and /tuáí/ oak tree, to be described. These forms are monosyllabic rather than bisyllabic, as witness the forms /doaiyáta/ he got well, and /tuaiyába/ on the oak tree. Although the final /i/ may be able to be traced back to some historical suffix, from the point of view of tone assignment, there is no justification for assuming that a boundary precedes the final /i/. If this is the case, then how is the tone contour of /doáí/ and /tuáí/ explained? One notices immediately that the high tone does not occur on the syllabic vowel, but only on the two contiguous glides. Since this is the case, and since in the set of rules already proposed to account for tones, one has rule (22), which states that a syllabic vowel becomes high toned if it precedes a high toned syllabic vowel, one might then propose that rule (22) be extended to state that any vowel becomes high toned before a high toned vowel which has the same value for the feature [syllabic] . That is, using the α notation to indicate this sameness of value, rule (22) can be reformulated as:

$$(35).\quad V \rightarrow +H \,/\, \underset{\alpha\,\text{syl}}{\underline{\quad}} \begin{bmatrix} V \\ +H \\ \alpha\,\text{syl} \end{bmatrix}$$

By utilizing the α notation, rule (22) has been generalized from the statement that a syllabic vowel preceding a high tone syllabic vowel becomes itself high toned, to the statement of (35), which says that whenever two vowels occur contiguously such that they are both glides or both syllabic, if the second is high toned, then so is the first. Thus /doáí/ would become /doaí/ by rule (21), and /doáí / by (35). /doaiyáta/, on the other hand, becomes /doaiyáta/ by rule (21) and (35) does not apply. Notice that for forms such as /maá/ <u>he gave</u>, where the specification for the vowel sequence is [+syl] [-syl] , rule (35) cannot apply. Notice also that the generalized form of (35) can only be stated if the first tone rule is given in the form of (21). If one had used the set of rules (17)-(20), or (31) and (19) and (20), where the hypothetical laryngeal /L/ was proposed, then the rules expressing the formation of the contour for /vúápai/ (rules (19) and (20)) could not be extended in any fashion to include the forms /doáí/ and /tuáí/, thus losing a significant generalization about the leftwards assimilation of tone when vowels agree in syllabicity.

The last point to be discussed concerns the HH contour on the second vowel cluster of a form. Examples of such forms include /ónai/ salt, /onáíši/ is it salt?, /onáíkɨdɨ/ with the salt, /oníába/ on the salt, /oníáva/ it is salt, /vúápai/ he brings, /vúápiava/ he does bring, /gasíúvikavoi/ brush, /sosóáñi/ they cried, /vakúánai/ to wash, /vovóítʸa/ aunt, /víótai/ to vomit, /víóíyata/ he vomited, /víói/ he vomits.[27] In the first set of examples, /ónai/ is the stem and /-ši/, /-kɨdɨ/, /-aba/ and /-ava/ are all suffixes. Earlier, it was mentioned that if a suffix beginning with a vowel were preceded by a stem which contained no consonant other than an initial consonant, a glide was inserted. An example of this is /víóíyata/, which consists of the stem /víói/ and the suffix /-ata/. However, if the stem does contain a consonant which is not the initial consonant, then the glide is not inserted. Rather, one of the stem vowels is deleted. Thus, one has /onai + ava/, yielding /oníáva/.

From all of the above examples, one sees that if the vowel in the second syllable is assigned the high tone by (21), then all of the following vowels in that cluster are high toned, giving a HH contour rather than an internal HL configuration. This assimilation occurs irrespective of whether the second member of the cluster is part of the stem or the initial vowel of a suffix. However, because of forms such as /áíyata/ _he caught up to him_, and /óíyata/ _he gave it_, where the stems /ái/ and /ói/ are bisyllabic and where the suffix has a prevocalic glide inserted by the rule for glide insertion, one sees that if the vowel following the second syllable vowel is preceded by a morpheme boundary and is a glide, then it does not become high toned.

Before stating the rule which will account for this fact, the motivation for a general glide insertion rule for suffixes beginning with a vowel when they occur after stems with no internal consonant will be given. When a stem of the shape (C)V(V)(V) is suffixed by, say, /-aba/, even if the final vowel of the stem is not /i/, there is no stem vowel deletion. Thus, from /moo/ + /-aba/, one has /móóaba/. If one assumes that at this point of the phonology that the form is actually /moo/ + /-yaba/, then one can state the conditions for stem vowel deletion quite simply as occurring in the environment ______ + [+syl] . The glide /y/ before the suffixes in /víóíyata/ and /móóyaba/ would then render these forms non-subject to the stem vowel deletion rule. Moreover, since the final shape of /moo + aba/ is /móóaba/ exactly as in /áíyata/, where the high tone does not carry over into the suffix, if the form of /mόόaba/ is assumed at this point of the phonology to be /moo + yaba/, then the following rule will state the condition under which one has internal high tone assimilation.

$$(36).\quad V \rightarrow +H \;/\; \begin{bmatrix} V \\ +H \end{bmatrix} \langle + \rangle_a \; \overline{\langle +syl \rangle_b} \; \begin{Bmatrix} V \\ C \end{Bmatrix}$$

In (36), the angle bracket notation is to be interpreted as in (21). (36) states that any vowel following a high toned vowel, if it is not word final, itself becomes high toned. If, however, the vowel in question is preceded by a morpheme boundary, then it cannot be a glide. If there is no morpheme boundary preceding the vowel in question, then the nature of the vowel with respect to the feature [syllabic] is irrelevant. Clearly, the rule for glide insertion must precede rule (36).

Notice that rule (36) does not apply to /mόόyaba/, as desired, if the glide insertion rule is ordered before (36). It will, however, correctly apply to /onai + aba/ to give /oníába/. We assume here that the stem vowel deletion rule is part of the glide insertion rule. A late rule, perhaps the last rule of the phonology, will then state the deletion of the glide /y/ when the stem does not end in /i/. Such a rule could be the following: [28]

(37). $y \rightarrow \emptyset \; / \left\{ \begin{matrix} [-\text{high}] \\ [-\text{front}] \end{matrix} \right\} + \_\_\_$

Rule (37) will yield /móóaba/ as desired. Rules (21), (29), (30), (35), and (36) account for the tonal phenomena of Northern Tepehuan. The following examples will show how these rules apply to different forms, given the specification for the feature [syllabic] for each of the vowels.

a.

/b	a	n	a	i /:
	+syl		+syl	-syl

/bánai/ coyote

bánai (rule 21)

b.

/d	a	a	k	a /:
	+syl	-syl		+syl

/daáka/ nose

daáka (rule 21)

c.

/v	u	a	p	a	i /:
	+syl	+syl		+syl	-syl

/vúápai/ he brings

vuápai (rule 21)
vúápai (rule 35)

d.

/	o	n	a	i +	š	i /:
	+syl		+syl	-syl		+syl

/onáíši/ is it salt?

onáiši (rule 21)
onáíši (rule 36)

e.

/t	a	i +	k	ɨ	d	ɨ /:
	+syl	-syl		+syl		+syl

/taíkɨdɨ/ with the fire

taikɨ́dɨ (rule 21)
taíkɨ́dɨ (rule 29)
taíkɨdɨ (rule 30)

f.

/v	i	o	i +	y	a	t	a /:
	+syl	+syl	-syl	-syl	+syl		+syl

/víóíyata/ he vomited

vióiyata (rule 21)
víóiyata (rule 35)
víóíyata (rule 36)

<u>2.3</u> A question which could be justifiably raised at this point is whether Northern Tepehuan is correctly described as having two tones and six tone sequences or combinations, or whether it is more correct to state that the tone system consists of six tones, two short, high and low, and four long, high level, low level, rising, and falling, after the fashion of tone descriptions of Oriental languages. The question of the "correct" designation is not a mere terminological quibble, rather, the issue at hand is the nature of the distinctive features of tone, the segment or segment groups to which they are assigned, and their relation to other phonological features.

It is evident from the data presented here, that the solution given for predicting the tones of Northern Tepehuan by rules is completely dependent on the assumption that long vowels are vowel sequences, and that the rising and falling tones are sequences of two pitches. Notice here that if one had used the features [rise] and [fall] to indicate contour tones rather than assuming that they were the sequences LH and HL, we could not have formulated rule (21) to account for the contours of bisyllabic forms. Instead, we would have had to state a completely different set of rules, none of which would be related. One might have the form:

(38). $V \rightarrow +H \ / \ C \underset{-long}{\underline{\quad}} \ CV(V)\#\#$

(38) would state that if a form is bisyllabic, and if the first syllable has a short vowel, then the tone of the first syllable is a short high level tone. Another rule would state that if the vowel is long in the first syllable, and the form is bisyllabic, then the first syllable has a rising tone. A third rule would have to account for the fact that if word were bisyllabic with no medial consonant, then the tone of the whole form is falling. It is not altogether clear how the last rule could be formulated, if at all, since the dynamic tone features have hitherto been thought to have a domain of only one syllable. Nevertheless, even if such a rule could be stated, the fact remains that the elegance and simplicity of the other solution are lost. More than that, however, to have to describe the tones in these terms denies that the tone configurations of bisyllabic words are formed by a single set of rules which ignore the length of the vowel in the first syllable. That is, our earlier rules simply stated that the high tone was located on the last vowel of the first syllable in a monosyllabic or bisyllabic word. If the first syllable contained a cluster, this would imply that one would phonetically perceive a rising tone. If there were no medial consonant, one

would perceive a falling contour. Aside from the fact that using [rise] and [fall] would obscure the uniformity of the nature of tone assignment in bisyllabic words, it is difficult to see how these features are to be used in describing contours such as those in /víói/ _he vomits_, or /doáí/ (as in /kašiš doáí/ _did he already get well?_), which are, respectively, a long high level tone followed by a fall, and a rising tone followed by a long high level tone. In the first case, one could say that the scope of the tone features was only the syllable, and that the bisyllabic word /víói/ was simply specified as high level on the first syllable and falling on the second. However, /doáí/ is not a bisyllabic word. How does one use these features to describe the tone configuration on this syllable?

The systems using the complex features [rise] and [fall] cannot describe these contours. However, since sequences of this type are somewhat rare, one might question the validity of these examples as a demonstration of the inadequacy of the [rise] and [fall] systems. If this were the only evidence, one might indeed question it. However, there is another set of languages whose tonal phenomena cannot be described by systems using [rise] and [fall] . These languages are the ones which make a phonemic distinction of five pitch levels.

They are rare, as is to be expected, being reported only in the Miao-Yao languages of Asia, and in the American Indian languages. Among those which have been reported as having five phonemic pitch levels are Trique,[29] a language of Mexico, Usila Chinantec,[30] another language of Mexico, and Ticuna,[31] a language spoken in Peru. The reports of these languages do not speak of contour tones *per se*, but rather of tone sequences occurring in single syllables. However, one must understand that this does not imply that there are no contour tones, this has simply been the standard manner of reporting contour tones in African and American Indian languages. The sequences are always described using numbers, or tone letters of the type proposed by Chao.[32] It is simple to understand why the contour tones were described in this manner, for the Indian languages in particular tend to have more than one rising tone and more than one falling tone. Thus, reference to a rising tone or a falling tone is ambiguous. The Trique dialect spoken in San Andres Chicahuaxtla, the one reported by Longacre, is perhaps an extreme example of a language with several rising and falling tones. The language is said to have thirteen of the possible twenty combinations of different pitch levels which can occur on a single syllable. Nine of these occur frequently. Even if the four of the thirteen contours which do not occur frequently could be collapsed so that they were lexically non-distinctive from the remaining nine, one still has the task of distinguishing between nine different simple contour tones. Notice that in Wang's system, only four simple contours can be described. What about the remaining five? This particular system can only describe four distinct simple contours because it only distinguishes between high and non-high rises and falls. Even if one could make the distinction in this system between say, a rising tone which rises from low to high as opposed to those rising from mid to high or from low to mid, one can still only have six distinctive simple contours. The reason for the inability is due to the fact that in order to describe five phonetic pitch heights using three features, one must define two of the pitch heights as combinations of these three features. Thus, the high-mid pitch is designated in Wang's system as $\begin{bmatrix} \text{+high} \\ \text{+central} \\ \text{-mid} \end{bmatrix}$. No rising tone which rises from, say, a mid-high pitch to a high pitch can then be designated, for the feature [rise] cannot order the beginning and end point of the range when one of the end points is itself defined by a set of features. That is, the feature specification of such a rising tone must be something of the sort, $\begin{bmatrix} +\text{high} \\ +\text{high} \\ +\text{central} \\ +\text{rise} \end{bmatrix}$, where somehow, the phonetic interpretative component can recognize that in this situation, $\begin{bmatrix} +\text{high} \\ +\text{central} \end{bmatrix}$ is to be realized as a simultaneous instruction, rather than as a sequential instruction. Alternatively, one could introduce a new height feature for the two intermediate levels and then restrict the system so that no two height features may be marked + in the same matrix unless the matrix also contains the specification [+contour] . Clearly we do not wish to resort to this alternative. The remaining alternative, however, is simply to deny that Trique has this tone system, which, obviously, is no solution.

Hence, from viewing these examples, and from the theoretical arguments previously stated, we conclude that a system using the features [rise] and [fall] is neither mechanically nor theoretically adequate to describe the tones of all languages. We thus abandon this approach and turn instead to a system which describes contour tones as sequences of pitch heights.

CHAPTER II: FOOTNOTES

1. In particular, see Li, K.P., "Tone perception experiment with appended test materials", POLA, no. 6, and Brotzman, R., "Progress report on Mandarin tone study", POLA, no. 8.

2. See Li, K.P., "Tone perception", and Wang and Li, "Tone 3 in Pekingese".

3. For example, see the spectrograms on pp. 122-123 in Denes, P.B., and Pinson, E.N., The Speech Chain.

4. There appears to be only one language which may be an exception to this statement, and that is a dialect of Mazatec about which Robert Longacre has spoken with me briefly. However, lacking any material on the language, I can make no further statement. Other apparent exceptions will be discussed in chapter IV.

5. I realize that there are segmental features which have been considered, which appear to be ordering features. One of these is the feature prenasalized . In SPE, pg. 317, there is some discussion of the status of this feature. I have nothing to add to that discussion. However, I think that the case presented here against the specific ordering

features [rise] and [fall] holds even if one must eventually permit the feature prenasalized , for I do not think that the two cases are exactly parallel.

6. Haj Ross has pointed out that another objection to treating vowel diphthongs in this manner is the fact that then the phonology would have to be complicated. That is, consider languages in which phonetic diphthongs arise from the deletion of intervocalic consonants. If one were to use features such as [rise] and [fall] to represent diphthongs, then the phonology must have in addition to a consonant deletion rule

(39). $C \rightarrow \emptyset \ / \ V_1 \_\_\_ V_2$,

the following:

(40). $V_1 \ V_2 \longrightarrow \begin{bmatrix} V_1 \\ \alpha \text{ rise} \\ \beta \text{ fall} \end{bmatrix}$ where α and β are determined by the nature of V_1 and V_2.

7. Kratochvíl, pp. 392-393.

8. There are those who may claim that the combination form is not low level, but rather, mid to low falling. In our spectrographic investigations, we have not found evidence to support this. In

any event, it is irrelevant to the discussion whether one specifies it as low level or mid to low falling.

9. Li, K.P., "Tone Perception".

10. This fact is graphically indicated by Chao in <u>Mandarin Primer</u>, where he represents the contour as  . (pg. 88)

11. In the following discussion, it is not crucial that my analysis of the third tone in isolation be accepted as , for the rules given can be changed to accomodate the configuration  . However, until evidence is presented to the contrary, we will continue to assume that the contour is  .

12. Wang, W. S.-Y., and Li, K.P., "Tone 3 in Pekingese".

13. Actually, the phonetic contour of these syllables is probably not a low to high rising tone, but a mid to high rising tone, for this sandhi contour has not been found to be readily distinguishable from underlying rising tones. We will discuss the adjustment rule in chapter III.

14. I am greatly indebted to Dr. Burt Bascom, not only for making the material about Northern

Tepehuan available, without which this part could never have been written, but also for his kind help and encouragement throughout the writing of this section.

15. Pike, Kenneth L., Barret, Ralph P., and Bascom, Burt, "Instrumental Collaboration on a Tepehuan (Uto-Aztecan) Pitch Problem", Bascom, Burton, "Tonomechanics of Northern Tepehuan", and Bascom, Burton, Proto-Tepiman.

16. Bascom, "Tonomechanics".

17. All examples are taken from the sources cited in 15 above. In these examples, all long vowels are represented as sequences of identical vowels. An accent mark on a vowel indicates that it is high toned, no mark, that it is low toned. Thus, the sequences VV, VV́, V́V, and V́V́, represent the long low tone, the long rising tone, the long falling tone, and the long high tone, respectively. An unmarked vowel followed by a marked vowel is therefore the sequence LH, or rising, as in /daáka/.

18. My thanks to Ken Hale for details on the morphological processes in the Piman languages.

19. The final /i/ is a glide which occurs with certain

verb forms. Glides will generally be transcribed exactly like the syllabic vowels unless they are morpheme initial. Although this may cause some confusion, this procedure has been chosen in order to cite the forms in exactly the same shape as they are found in Bascom's works.

20. I am indebted to Ken Hale for this historical information.

21. For a discussion of the feature [syllabic], see SPE, pg. 354.

22. This use of the angle brackets differs slightly from that in SPE, in particular, as discussed on pp. 347-348, and pp. 394-395. There the angle bracket notation is used as an abbreviation for an "if and only if". The interpretation "if a then b" allows the expansion to include one more string than the "if and only if" interpretation. Obviously, one could write a set of rules which did not make any use of the angle bracket notation which gives the same results as rule (21). However, I feel that (21) is the correct formulation.

23. Actually, the environment of (21) will have to be modified slightly to permit consonant clusters

to occur after the vowel in question. Thus, the environment of (21) must be:

$$ {}_{\text{Stem}}\Big[\, C_o (V) \left(\begin{bmatrix} V \\ -\text{syl} \end{bmatrix} \right)_o^2 C_o \; \underset{\langle +\text{syl} \rangle_a}{\underline{\qquad\quad}} \; \left\langle C_o^2 \begin{bmatrix} V \\ +\text{syl} \end{bmatrix} \right\rangle_b $$

24. For all forms, if one holds the view that clusters are formed by the deletion of a vowel, then the vowel of the penultimate syllable can never be deleted unless it is the vowel /i/. Because /i/ causes surrounding consonants to become palatalized, if the forms are all entered with a CVCVCV... shape, then the palatalization of certain consonants becomes predictable.

25. In order to save the laryngeal solution, one might think that the glide insertion rule could be changed so that it inserted a consonant instead of a glide, and then ordered before the tone rules. However, if the glides are inserted before rule (21), then there are problems with items such as /mói/ and /ái/, where the underlying forms were to be /moLo/ and /aLi/. These forms are subject to the glide insertion rule. (The reasons for stating that a glide is inserted in the form /moo + aba/ is given on pg. 124.) However, the conditions for inserting the glide are simply that the stem contain no internal non-vocalic,

unvoiced segment, or in other words, be of the form, (C)VV(V). If the hypothetical laryngeal exists in /móo/ and /ái/ when the glide insertion rule applies, then it must be an exception to the conditions of the glide insertion rule. Notice that in order to make rule (31) work, the laryngeal must be present when the glide insertion rule applies. However, then the hypothetical laryngeal is a complete exception to this rule. It would not only be behaving differently from all of the real consonants of Northern Tepehuan, but also from the laryngeal which does exist in Northern Tepehuan, /h/. That is, one finds the forms /ááhɨava/ *he does catch up to him*, and /vúápiava/ *he does bring*, which consist of the stem /ááhɨi/ followed by /-ava/, and the stem /vúápai/ followed by /-ava/. In neither of these cases is the glide inserted. Thus, in maintaining this position, one must not only claim the existence of an underlying phoneme which has no phonetic reality, but also one which, at least for one rule of the phonology, must behave as a complete exception.

26. A detailed discussion of the theoretical implications of positing underlying phonemes which have no phonetic reality is given in a paper by Paul

Kiparsky, "How Abstract is Phonology?".

27. Obviously, from these forms we must change the environment of (21) again. The complete environment is:

$$\text{Stem}\left[C_o(V)\left(\begin{bmatrix}V\\-\text{syl}\end{bmatrix}\right)_o^2 C_o \; \overline{\langle +\text{syl}\rangle_a} \; \left\langle (V)\; C_o^2 \begin{bmatrix}V\\+\text{syl}\end{bmatrix}\right\rangle_b\right.$$

where the second syllable can contain a vowel cluster. This does not affect the application of the rule in any way, of course.

28. The feature [high] used in rule (37) is not to be confused with the feature [high tone] . The former refers to tongue height, the latter to pitch height.

29. Longacre, R.E., "Five Phonemic Pitch Levels in Trique".

30. Skinner, L.E., "Usila Chinantec Syllable Structure".

31. Anderson, L., "Ticuna Vowels: with Special Regard to the System of Five Tonemes".

32. Chao, Y.R., "A system of tone letters".

CHAPTER III: THE DISTINCTIVE FEATURES OF TONE: A NEW PROPOSAL

3.0 In chapter II, we showed that there are many inadequacies inherent to a system of features for tone which ignores the relationship between sonorant sequences and contour tones, and describes contour tones by means of such features as [rise] and [fall] . We therefore propose that the distinctive features of tone are features of pitch height, and that contour tones are represented as sequences of these features, each one of which is uniquely associated with some sonorant segment. We shall now state what these features are, and how they are to be used. The phonetic correlates to the features proposed will be given in chapter IV.

Since there are languages which have been reported to have five distinctive pitch levels, it is obvious that there must be at least three pitch features. It does not follow, however, that the three pitch features must exist on a par, that is, that there is not a hierarchy of pitch features.

Note that the majority of languages have either two or three contrastive pitch levels, that is, high versus non-high, or low versus non-low, or high, low, and mid (non-high, non-low). This fact would seem to indicate that the primary distinctions for which we must account are high versus non-high, and low versus non-low. Let us therefore define $[-F_i]$, for all pitch features $[F_i]$, as referring to a rest position. It then follows that a matrix specified as $\begin{bmatrix}\text{-high tone} \\ \text{-low tone}\end{bmatrix}$ will be phonetically realized as a mid tone.

This partitioning of pitch features into two main oppositions is clearly analogous to the partitioning of vowels with respect to tongue height by means of the features [high] and [low] . We notice in particular, that in none of the languages which have two or three contrastive pitch heights, examples of which are Bambara, Northern Tepehuan, and Standard Japanese with two contrastive pitch heights, Cantonese, Ewe, and Standard Thai with three contrastive pitch heights, is the so-called "mid" tone distinct from the non-high, non-low tone. Thus, although there are languages which are reported to have four and five contrastive pitch levels, we do not feel that the introduction of a pitch height feature [mid] to account for the possibility of five distinctive levels is justified. Rather, we feel that the intermediate levels, those traditionally described as high-mid and low-mid, or in Chao's "tone letter" notation, levels 4 and 2 respectively, are the result of a secondary articulatory mechanism imposed on the primary mechanisms. The primary features are thus [high tone] and [low tone] .[1] These features have the obvious phonetic correlates to high and low pitch. The secondary feature which gives rise to the intermediate levels is the feature [modify] .

There are several reasons which motivate us to choose the set of features [high tone] , [low tone] , and [modify] , as opposed to the set [high tone] , [low tone] , and [mid] . The first of these is simply the fact previously stated, that the mid tone in languages with no more than three contrastive pitch heights is always identical to the one which is also designated as $\begin{bmatrix}\text{-high tone}\\ \text{-low tone}\end{bmatrix}$. This redundancy can, of course, be expressed by the rule: $\begin{bmatrix}\text{-high tone}\\ \text{-low tone}\end{bmatrix}$ $\begin{bmatrix}\text{+mid}\end{bmatrix}$.
However, to have to state such a redundancy explicitly is to deny the implicit statement that a pitch which is neither high nor low is obviously mid. More important, however, is the fact that if the feature [mid] is on a par with the features [high tone] and [low tone] , then we do not have the ability to recognize that languages with more than three contrastive pitch heights are highly marked and therefore rare. With the set of features we propose, this would be reflected by the fact

that these languages would have to make distinctive use of a secondary feature which is normally non-distinctive.

It is difficult to see what the articulatory correlate to the feature specification [+mid] could be, other than the redundant one also given by the feature matrix $\begin{bmatrix}\text{-high tone}\\ \text{-low tone}\end{bmatrix}$. If the feature [mid] were used with this interpretation, then we cannot see how the matrix for the mid-high pitch, for example, can have any interpretation. That is, the mid-high pitch would have to be specified as $\begin{bmatrix}\text{+high tone}\\ \text{+mid}\end{bmatrix}$. However, [+mid] , on the one hand, is equivalent to $\begin{bmatrix}\text{-high tone}\\ \text{-low tone}\end{bmatrix}$, and [+high tone] , on the other hand, must be interpreted as [+high tone] . Thus, there seems to be a conflict in the phonetic instructions represented by that matrix using [mid] to define the high-mid pitch level. Such a conflict will not arise with the feature [modify] , for [+modify] is associated with a phonetic interpretation distinct from that given to the matrix $\begin{bmatrix}\text{-high tone}\\ \text{-low tone}\end{bmatrix}$. We will go into detail about the articulatory correlate to this feature in the next chapter.

<u>3.1</u> The features thus proposed are: [high tone] , [low tone] , and [modify] . From the discussion in

chapter I, there is obviously a fourth prosodic feature, [stress] , but we shall not discuss it here. The following table will demonstrate how the different pitch contours are specified with respect to these features. Each column matrix is understood to be part of the feature matrix of some sonorant segment. Since three binary features, even when restricted such that only five of the possible eight combinations can occur in any given matrix, can produce twenty-five different sequences of two matrices, one hundred and twenty-five different sequences of three, and so forth, we shall not present an exhaustive table. The restrictions on the features will be stated after the table has been given, since they will then be more comprehensible.

<u>TABLE 3.1</u>

	˥	˦	˧	˨	˩	˥˥	˩˩
high tone	+	+	-	-	-	+ +	- -
low tone	-	-	-	+	+	- -	+ +
modify	-	+	-	+	-	- -	- -

	˧˧	˧˥	˩˥	˦˥	˩˨	˨˧
high tone	- -	- +	- +	+ +	- -	- -
low tone	- -	- -	+ -	- -	+ +	+ -
modify	- -	- -	- -	+ -	- +	+ -

	˧˦	˥˧	˧˩	˥˦	˥˩
high tone	- +	+ -	- -	+ +	+ -
low tone	- -	- -	- +	- -	- +
modify	- +	- -	- -	- +	- -

	˧˩˧	˩˩˧
high tone	- - -	- - -
low tone	- + -	+ + -
modify	- - -	- - -

The redundancy rules for these features are the following:

(1). $[+\text{high tone}] \rightarrow [-\text{low tone}]$

(2). $[+\text{low tone}] \rightarrow [-\text{high tone}]$

(3). $\begin{bmatrix}-\text{high tone}\\ -\text{low tone}\end{bmatrix} \rightarrow [-\text{modify}]$

Rules (1) and (2) state the obvious restrictions, namely, that a pitch cannot be simultaneously high and low. Rule (3) states that the secondary feature [modify] cannot be specified as + unless the matrix is also specified as either [+high tone] or [+low tone] . We believe that the features [+high tone] and [+low tone] represent particular positions of the glottis, namely, heightened glottis and lowered glottis, respectively. The neutral,or mid position, we therefore designate by the feature specification $\begin{bmatrix}-\text{high tone}\\ -\text{low tone}\end{bmatrix}$. We shall show when we present the marking conventions that this mid position is the unmarked position. Rule (3) thus states that the modifying mechanism can only be effected when the glottis is not at the neutral position. We must state that we only hold this position tentatively, for we have no strong evidence to support this view. We feel, however, that convincing justification for accepting or rejecting this view can only come through more intensive investigation of the articulatory movements involved in producing the intermediate pitch levels in the five level languages, and of the movements involved in the "up-stepping" languages, of which latter type, Acatlan Mixtec is reported to be one.[2] Until such studies have been made, we shall accept (3) as correct.

From Table 3.1, we see that short vowels may only carry level contours. Long vowels and other sonorant clusters, however, may have either level or contour tones. The mid tone is denoted by the feature matrix,

$$\begin{bmatrix} \text{-high tone} \\ \text{-low tone} \\ \text{-modify} \end{bmatrix} .$$

3.1.1 Let us now consider again the tone sandhi in Mandarin and the lexical representation of its tone system in terms of the features just presented. The question which we must answer before we can discuss the sandhi is simply whether Mandarin has four lexical tone sequences or five. Rephrased, does Mandarin have only the four underlying forms of the citation tones in the lexicon with the "neutral tone" syllables derived from these, or do the neutral toned syllables of lexical polysyllabic items have a distinct tonal representation? Notice that the latter alternative does not deny that some neutral tone syllables may be derived from lexical items with one of the four underlying citation contours by means of stress placement, it simply questions the nature of the lexical specification of the unstressed syllables in such items as the bisyllabic words presented in Table 2.2. There are several hundred such bisyllabic lexical items with the phonetic stress pattern, "stressed - unstressed", in which the tone of the second syllable must somehow be designated as "neutral" in order to yield both the correct stress pattern as well as the correct tone pattern. Cheng Chin-chuan, in his thesis, Mandarin Phonology, argues convincingly that the unstressed and neutral toned quality of these syllables cannot be predicted in any way from the internal syntactic structure of the items.[3] Thus, he argues that the neutral toned quality of these syllables must somehow be denoted, from which fact the [-stress] quality of the syllables will be predictable. However, unlike the more traditional analyses of Mandarin such as those proposed by Hockett[4] or Martin[5], in which the neutral tone is considered as a fifth distinct lexical tone contour, Cheng argues that there should only be four distinct contours in the lexicon and that the neutral tone should always be derived from these. We disagree with Cheng and feel that the traditional analysis, that the lexical neutral tone syllables are specified with a fifth pitch contour, is correct.[6] Let us assume temporarily that this fifth contour is $\begin{bmatrix}\text{-highT}\\\text{-lowT}\end{bmatrix}\begin{bmatrix}\text{-highT}\\\text{-lowT}\end{bmatrix}$. We will discuss the reasons for choosing this configuration later.

In chapter II, section 2.1.5, we gave a number of reasons for considering the lexical form of the Mandarin third tone syllables as having the lexical pitch specification, [+lowT] [+lowT] . The high level tone syllables are specified as [+highT] [+highT] . The rising tone syllables are specified as $\begin{bmatrix}\text{-highT}\\\text{-lowT}\end{bmatrix}$ [+highT] , or, in other words, as a mid to high rising tone. We feel that this specification, which is in accordance with the traditional descriptions of this tone, is more correct than an underlying specification [+lowT] [+highT] . In fact, since the latter is what we claim is the sandhi form of the Mandarin third tone when it occurs before another third tone, the different starting points of the rise may be what accounts for the difference some people claim exists between the underlying rising contour and the one resulting from the sandhi rule applying to the low level tone. We are not convinced, however, that this difference is actually present, since identification of those sandhi created forms as distinct from underlying forms with rising tones cannot generally be made consistently by native speakers.[7] Moreover, the times when the identification can be made, that the underlying form of the syllable is low level, cannot be clearly attributed to only a difference in the starting point of the rise. There are probably many other linguistic and extra-linguistic factors involved in the identification of these syllables as arising from underlying low level tone syllables. Thus, although [+lowT][+highT] is the sandhi form given by the rules we presented in chapter II, we think that there is probably another rule in the phonology which adjusts the starting point of the rising tone resulting from sandhi to $\begin{bmatrix}\text{-highT}\\\text{-lowT}\end{bmatrix}$, thus rendering these syllables phonetically indistinguishable from the underlying rising tone syllables. The syllables with an underlying falling tone will be specified as [+highT][+lowT]. If only these four sequences were distinctive, then one or the other of the two features [high tone] or [low tone] would not have to be used distinctively, that is, one would not have to specify any + values for one or the other of these two features. However, since we claim that there are five underlying sequences, then both of these features must be used distinctively, in order to get five different combinations of two pitch sequences. The question which now arises is what is the specification of the neutral tone syllables in the lexicon?

Before discussing the answer to this question, recall that the segmental structure of the unstressed, or neutral tone, syllables in the bisyllabic lexical items must have the same syllable shape as any other Mandarin syllable, namely, $C_o(G)\begin{bmatrix} V \\ +syl \end{bmatrix}\begin{bmatrix} +son \\ -syl \end{bmatrix}$.

The reason for this comes from two observations. The first is that often, when the truncation rule for unstressed syllables is applied to a syllable closed by a nasal consonant, the preceding vowel is nasalized. Thus, were we not to assume that the second syllable in an item such as /day shanq/ 戴上 to put on (clothes or ornaments),[8] was entered in the lexicon as /šaŋ/, we would have to state that the nasal vowels were distinctive in Mandarin, but only in unstressed syllables. We do not wish to make such a claim. The second observation comes from the fact that apparently for some speakers and with some lexical items, the truncation rule deleting the final sonorant in unstressed syllables is not effected. In those instances, the syllable is not quite as short as otherwise, and the entire vocalic nucleus, pure long vowel, diphthong, or vowel plus nasal, is retained. In these instances, the tone of the entire sonorant cluster is precisely that predicted from the tone contour of the preceding syllable, that is, some long level tone. It is rather difficult to obtain material for laboratory studies of these non-truncated unstressed syllables, for we have found that most speakers in reading bisyllabic items either recognize it immediately as one with a "stressed - unstressed" pattern, and thus read the second syllable as a very short unstressed syllable, deleting the final sonorant, or they will only vaguely recognize the form. In the latter case, this usually means that the word is not normally used in their speech and that they will articulate it as two stressed syllables. We have, in fact, found only one instance with one speaker where the syllable could be definitely stated to be unstressed, as reflected by the intensity, and where the truncation rule did not apparently apply. Many speakers will claim that they can hear full diphthongs in <u>short</u>, unstressed syllables, thus posing a potential counter argument to our claim about the length distinction in vowels in Mandarin. However, in our investigations, we have not found any spectrographic evidence showing that diphthongs exist as the vocalic nucleus in short unstressed syllables. That is, the vowel formants in the short syllables were consistently level, rather than bending, as would have been the case if there were diphthongs. We feel that the perception of diphthongs might arise from two sources; 1) a knowledge of the underlying form of the syllable and the imposition of this knowledge on the perception of the syllable, and 2) a confusion of the term "unstressed". We do not feel that the latter source of confusion is widespread, but one does find occasional references in the literature to phonetically non-primary stressed syllables as unstressed. This occurs, for example, in Kratochvíl's work. He calls the bisyllabic pairs such as /jann juh/ 站住 <u>to stand firmly</u>, which is characterized by the stress configuration 2 1 , a sequence "unstressed syllable - stressed syllable".[9] At least in this one case, secondary, or possibly tertiary, stress has been confused with [-stress] .

Thus, the lexical syllables which are phonetically unstressed will, like all other lexical syllables, have the structure $C_o(G) \begin{bmatrix} V \\ +syl \end{bmatrix} \begin{bmatrix} +son \\ -syl \end{bmatrix}$. The pitch specification of these syllables will be $\begin{bmatrix} -highT \\ -lowT \end{bmatrix} \begin{bmatrix} -highT \\ -lowT \end{bmatrix}$, or a sequence of mid tones. We have chosen this specification for two reasons. The first is that it is precisely that one which follows stressed syllables with the pitch contour [+lowT] [+lowT] , and therefore is a contour which does occur phonetically. The second is that we feel that $\begin{bmatrix} -highT \\ -lowT \end{bmatrix}$ is the "unmarked" specification of pitch features, and that, therefore, it is characteristically the lexical specification of the pitch features of any sonorant segment whose pitch is predictable in any manner. Hence, it would be the lexical specification of all sonorants in English, for example, since high tone in English occurs only with stress and stress is predictable. The Mandarin tones and their feature specifications are the following: (each column matrix is presumed to be part of the feature matrix of some sonorant segment)

<u>TABLE 3.2: MANDARIN TONES</u>

	(1) ˥	(2) ˧˥	(3) ˩	(4) ˥˩	(5) ˧
high tone	+ +	- +	- -	+ -	- -
low tone	- -	- -	+ +	- +	- -
modify	- -	- -	- -	- -	- -

Let us now consider the sandhi and neutral tone realization rules. We assume that before either of these rules is applied, stress is assigned. We shall not be concerned with the cyclic stress assignment rules, nor the stress deletion convention for phrases such as verb phrases with directional adverbs. We shall simply assume that by the time the sandhi and neutral tone realization rules apply, the proper stresses are specified. However, in addition to these stress rules, we must add one more, and that is the rule assigning [-stress] to the second syllable of bisyllabic lexical items with the phonetic stress pattern "stressed - unstressed". An example is the word /dong shi/ 東 西 thing. For this and the other lexical items whose stress pattern cannot be determined from their syntactic structure, there is the following stress rule.

$$(4).\quad V \rightarrow [\text{-stress}] / \begin{bmatrix} \underline{\quad\quad} \\ \text{-highT} \\ \text{-lowT} \end{bmatrix} \begin{bmatrix} \text{+son} \\ \text{-highT} \\ \text{-lowT} \end{bmatrix}$$

3.1.2 The tone sandhi rule must follow the stress rules, for with the exception of the lexical item stress rule, rule (4), stress assignment is cyclic and the sandhi rule is not. The sandhi rule, which changes a low level tone to a rising tone whenever it occurs before another low level tone, cannot be cyclic because it must apply to any continuous sequence of low level tones in a phonological phrase, irrespective of the internal bracketing of the phrase. Hence, given an expression such as /woo mae bii/我 買 筆 I buy a pen, where each of the syllables is specified in the lexicon with a low level tone, the rule must apply simultaneously to the first two syllables. The domain of the rule is not well defined, but must be stated as "within a breath group or phonological phrase". That is, when there is a strong pause between two items with lexical low tones in any given utterance, the first syllable in this instance will not be subject to the sandhi rule, but rather to the phrase final lengthening rule affecting low level tones as discussed in chapter II. The sandhi rule is thus expressed as:

$$(5).\quad \begin{bmatrix}+\text{son}\\+\text{lowT}\end{bmatrix} \rightarrow [+\text{highT}] \,/\, \begin{bmatrix}\text{V}\\+\text{lowT}\end{bmatrix} \_\_\_ (\#\#)\text{C}_{\text{o}}(\text{G}) \begin{bmatrix}\text{V}\\+\text{lowT}\end{bmatrix}$$

The effect of (5) is to change every instance of a low level tone, the sequence $[+\text{lowT}]\ [+\text{lowT}]$, to a rising tone, the sequence, $[+\text{lowT}][+\text{highT}]$, whenever it occurs before another low level tone syllable inside of a phonological phrase. As has already been discussed in chapter II, syllables with low level tones which occur phrase finally, and hence, also in isolation, have an extra long form and a rising tonal contour toward the end of the utterance of the syllable. The length was given by the rule:

$$(6).\quad \emptyset \rightarrow \begin{bmatrix}\text{V}\\-\text{syl}\\+\text{lowT}\\\alpha\text{F}\end{bmatrix} \Big/ \left[\begin{bmatrix}\text{V}\\+\text{lowT}\\\alpha\text{F}\\+\text{stress}\end{bmatrix} \_\_\_ \begin{bmatrix}+\text{son}\\+\text{lowT}\end{bmatrix}\right]_{\text{phon. phrase}}$$

Notice that this rule only applies when the syllable is in absolute final position, for if there is any intervening syllable, stressed or unstressed, these syllables do not become lengthened. This, one will recall, was the primary justification given for considering the lexical forms of these syllables to contain only the normal long vowel, diphthong, or vowel plus nasal cluster, and for considering the underlying pitch specification to be low level. The tone configuration of the extra-long form is then adjusted by a rule which also determines the pitch configuration of unstressed syllables following low toned syllables. This latter rule, which we shall discuss in detail shortly, and whose precise form we will give later, will assign a mid tone, or $\begin{bmatrix}-\text{highT}\\ -\text{lowT}\end{bmatrix}$, to the last sonorant of the extra-long forms, thus giving the shape $[+\text{lowT}][+\text{lowT}]\begin{bmatrix}-\text{highT}\\ -\text{lowT}\end{bmatrix}$. Applying rules (5) and (6) and the unstated final sonorant adjustment rule to the example given previously, we obtain the following derivation: (We assume that the specification of /woo mae bii/ before these rules apply is the one given in (7) below.)

```
(7).  #w   ɔ     ɔ     ##m   a     i    ##b   i     i     #
          +lowT +lowT       +lowT +lowT      +lowT +lowT
```

applying rule (5) we get:

```
(8).  #w   ɔ     ɔ      ##m   a     i      ##b   i     i     #
          +lowT +highT       +lowT +highT       +lowT +lowT
```

applying rule (6) we get:

```
(9).  #w   ɔ     ɔ      ##m   a     i      ##b   i     i     i     #
          +lowT +highT       +lowT +highT       +lowT +lowT +lowT
```

applying the adjustment rule for the starting point of the sandhi forms mentioned on pg. 151, we get:

```
(10). #w  ɔ       ɔ     ##m  a       i    ##b  i      i      i    #
     -highT  +highT   -highT  +highT   +lowT  +lowT  +lowT
     -lowT            -lowT
```

finally, applying the last sonorant adjustment rule, we get:

```
(11). #w  ɔ       ɔ     ##m  a       i    ##b  i      i      i    #
     -highT  +highT   -highT  +highT   +lowT  +lowT  -highT
     -lowT            -lowT                          -lowT
```

3.1.3 The exact nature of the neutral tone after each of the citation tones has been a matter of some dispute among linguists. Chao, for example, states that after the high level tone, an unstressed syllable is low pitched, after the rising tone, it is mid pitched, after the falling tone it is low pitched, and after the low level tone it is mid, or mid-high, pitched.[10] His statement, however, is not entirely supported by the acoustic findings of several experiments. Wang and Li,[11] for example state that they have found the unstressed syllables to be high toned after the high level tone syllables and after the rising tone syllables. Dreher and Lee[12] report essentially the same results. We shall discuss the falling quality which they noticed in their tests in chapter IV. Suffice it to say that we feel that not only are the various contours they describe non-distinctive, but also that they are the natural phonetic results of the intensity contour of these syllables. The question which is of some interest, and about the answer to which we shall only briefly speculate, is why did Chao state that the unstressed syllables following syllables ending in a high pitch to be low? The answer may be that the relatively low intensity of these syllables became equated with low pitch. Also, the fact that the intensity of these syllables sharply decreases, leading, therefore, to a generally falling quality of the pitch, may have led him to feel that the low point was the one to be described. We feel, however, that since the starting pitch of the unstressed syllables following the high tone or rising tone syllables has consistently been noted as high, as opposed to those following the other syllables, that this is the property which we must reflect. The falling quality will be present if and only if the speaker articulates the syllables with a rapidly decreasing intensity. We notice, in support of this statement, the contrast between the staccato, falling nature of the unstressed syllable when it is phrase final as opposed to its nature when it is not phrase final but is followed either by more unstressed syllables or is medial between two stressed syllables. In the first case, the pitch contour is usually level since the intensity remains a constant on non-final unstressed syllables before unstressed syllables. In the second case, the contours may be level or may be affected by the starting point of the following stressed syllable.

Hence, in agreement with Wang and Li, and Dreher and Lee, we state that the unstressed syllables after syllables ending in a high pitch are high pitched, those after the falling tone are low pitched, and those after the low toned syllables are mid pitched. The rule stating the pitch value of the unstressed syllable after high toned, rising toned, and falling toned syllables is the following:

$$(12).\quad V \rightarrow \begin{bmatrix} \alpha\,\text{highT} \\ \beta\,\text{lowT} \end{bmatrix} / \begin{bmatrix} \text{V} \\ \text{-lowT} \\ \text{+stress} \end{bmatrix} \begin{bmatrix} \text{+son} \\ \alpha\,\text{highT} \\ \beta\,\text{lowT} \end{bmatrix} \text{(C)(G)} \underline{\qquad\quad} \atop \text{-stress}$$

As one readily sees, rule (12) simply states that in the case of the three contours which are non-low in their first segment, the neutral tone is an assimilation to the second pitch value of the preceding syllable. Notice, however, that in order to implement this rule, we must assume that the sandhi form of the low level tone syllables has already been adjusted from $\begin{bmatrix}\text{+lowT}\end{bmatrix} \begin{bmatrix}\text{+highT}\end{bmatrix}$ to $\begin{bmatrix}\text{-lowT} \\ \text{-highT}\end{bmatrix} \begin{bmatrix}\text{+highT}\end{bmatrix}$, or, in other words, that these sandhi forms are, in fact, featurewise indistinguishable from the underlying rising tones. We have already stated that this rule exists in the phonology and now we state that it follows the sandhi rule and precedes the neutral tone realization rule. This means, of course, that in the derivation given for /woo mae bii/, line (10) <u>must</u> precede line (11). The second part of the neutral tone realization rule must handle both the cases of the tone of the final sonorant in phrase final extra-long syllables as well as unstressed syllables following a low level tone.

Notice that in the environment of rule (12), no mention is made of ## (word boundary). Because unstressed syllables are clitics to the preceding stressed syllables, we feel that the ## boundaries which are usually used to delimit a phonological word do not exist here, since these are not phonologically two words, but one. Thus the environment of rule (5) must only contain optional word boundaries, for it applies both within a phonological word as well as across word boundaries.[13] Since some of the unstressed syllables can be part of a syntactically formed phrase, an adjustment rule must erase the word boundaries before (12) applies. The rule will simply delete the word boundary before any [-stress] syllable which occurs in a polysyllabic word, be it lexical or derived.

Thus, in the verb phrase /kai jinn lai/, which consists of the main verb /kai/ followed by the complex adverbial /jinn lai/, the rule will apply to the string /##kai##jinn##lai##/ to give /##kai jinn lai##/, since the last two syllables are clitics. The rules for assigning the stress in this phrase were discussed in chapter I. We feel that the boundary erasure rule is well motivated, for it is generally thought that ## do not exist between clitics and the item to which they are attached. Thus, it may in fact be the case that this rule is not just a Mandarin rule, but is part of some universal convention concerning clitics.

Before we can state the second part of the neutral tone realization rule, we must first recall the early rule of the phonology which assigns [+stress]. If stress is interpreted as intensity in Mandarin, which, in fact, our laboratory studies show it to be, then clearly only [+syllabic] segments can be specified as [+stress]. This is, roughly speaking, the assumption which underlies most discussions of what is commonly recognized as stress. It follows then, that every [-syllabic] segment should automatically be [-stress]. All rules dealing with [+stress] assignment can therefore be assumed to concern only syllabic vowels. The non-syllabic sonorants of all syllables, including stressed ones, can be assumed to be specified as [-stress] at the point at which the neutral tone realization rules apply. We can therefore state the remaining part of this rule as the following:

$$(13).\quad [+son] \rightarrow \begin{bmatrix}-highT\\-lowT\end{bmatrix} / \begin{bmatrix}V\\+lowT\\+stress\end{bmatrix} [+son]\ C_o(G)\underline{\qquad}_{-stress}$$

Notice that (13) applies to both the final sonorant in the extra long forms arising from the application of rule (6), as well as to the vowel of clitics. All of the extra long syllables are obviously [+stress] because rule (6), which gives the extra length, only applies to [+stress] syllables.

What about the second sonorant in these unstressed syllables, and what about sequences of clitics such as in phrases like /kai jinn lai/, where both the second and third syllables are clitics? The tone of the entire unstressed syllable, in the event that there is just one, must be specified as level, for if the truncation rule does not apply, the syllable will be long and the pitch contour level. In the case of several clitics, each has the same pitch specification as the first. For example, in /kai jinn lai/, the last two syllables are clitics, although they are not lexical neutral tone syllables, i.e. they are not specified with contour (5) in table 3.2 in the lexicon, but instead, are assigned [-stress] by the stress rules for verb phrases with directional complements discussed in Chapter I. Here, the phonetic pitch specification of <u>all</u> the sonorants in the two unstressed syllables is [+highT] . Rule (13) only takes care of the first one. Utilizing the fact that all the segments in these clitics are specified as [-stress] , we can state the following rule to account for the pitch specification of the remaining sonorants.

(14). $$[+\text{son}] \rightarrow \begin{bmatrix} \alpha\,\text{highT} \\ \beta\,\text{lowT} \end{bmatrix} / \begin{bmatrix} \text{V} \\ +\text{stress} \end{bmatrix} [+\text{son}]\ C_o(G) \begin{bmatrix} \text{V} \\ -\text{stress} \\ \alpha\,\text{highT} \\ \beta\,\text{lowT} \end{bmatrix} X \ \overline{}\ -\text{stress}$$

where X is the schemata abbreviated by

$([+\text{son}]\ C_o(G)(V))^*$

Rule (14) simply states that all unstressed sonorants following the first unstressed sonorant are assigned the same pitch specification as the first one. Thus, <u>X</u> in the rule simply represents any string of segments which contains no word boundaries (##). Rule (14) therefore applies to all of the sonorants in /kai jinn lai/ after the /i/ in /jinn/.

Notice that one peculiar effect of (14) would be to assign distinctive pitch levels to all sonorants following the first unstressed one, including morpheme initial nasals and prevocalic glides, which otherwise do not carry distinctive tone. That is, word initially we do not feel that the initial [-syllabic] sonorants carry any significant part of the pitch contour, but rather are "neutral" with respect to the tone of the syllable. In the case of clitics, however, the nasals and prevocalic glides become word medial. Hence, we feel that they are involved in carrying the pitch contour of the word. This phenomenon is particularly noticeable in rapid speech in the articulation of a phrase such as /ta men/ 他們 *they (or them)*, where the entire syllabic nucleus of the unstressed second syllable may be deleted, leaving only /tam/. In this case, the morpheme initial nasal certainly takes on the tone of the preceding vowel.

One could state (14) not as a separate rule, but as part of the neutral tone realization rule, at least for the first part of the realization rule, rule (12). However, because we cannot see how this can be done for the second part, rule (13), and because we are not convinced that it is one process even if we could formulate the assimilation as part of the second half of the realization rule, we shall leave the formulation as stated. These, then, are essentially the tone rules for Mandarin. There are some minor adjustment rules, such as the "leveling" of a rising tone when it is between two high toned syllables, but we shall not discuss them here. Following the realization rules, (12) and (13), and the assimilation rule (14), of course, is the unstressed syllable sonorant deletion rule, which deletes the sonorant following the syllabic in unstressed syllables. If the sonorant to be deleted is a nasal, there is a nasalization rule affecting the preceding rule.

3.2 In all of the preceding discussion, we have tacitly assumed that the segments whose pitch features were of interest were those composing the syllabic nucleus of any syllable. However, if our claim that the features presented here are the features of pitch is to be upheld, we must also indicate the values that these features can assume for segments other than those in the syllabic nucleus. In addition, it is interesting to inquire if there are marking conventions to support the claim made of the feature specification of lexical neutral tone syllables, that each member of the syllabic nucleus of each of these syllables is specified as $\begin{bmatrix}\text{-highT} \\ \text{-lowT}\end{bmatrix}$, and that this is in fact the "unmarked" specification of these features. We shall therefore now discuss the specification of these features for various classes of segments and then present a set of marking conventions to express the constraints and the optimal configurations.

3.2.1 It has been a generally noted fact that only the syllabic nucleus can be articulated with, or perceived clearly to carry, a pitch contour. Stop consonants, although sometimes noted to have a characteristic pitch, have never been known to carry different pitches without having the pitch difference also reflected in a parallel manner by a voicing difference. That is, although voiced consonants have been said to have a "lower" pitch than voiceless ones, in particular, with respect to their effect on contiguous vowels, one does not find both high and low pitched voiced obstruents. We believe that this is not an accident, but rather the natural consequence of the vocal tract configuration in the articulation of non-sonorants. That is, the inherently higher pitch of the voiceless non-sonorants may be a direct consequence of the spread, raised, and hence, probably tenser, vocal cord configuration, as opposed to the closed, lowered, and hence, probably laxer, vocal cord configuration of voiced non-sonorants. However, even if the vocal cord tension in the voiced non-sonorants could be variably adjusted, as it can be in the articulation of voiced sonorants, where the vocal cords are also closed, the pitch variation which results in vowels probably does not arise in the case of non-sonorants due to the lack of an unobstructed vocal tract. That is, the requirements necessary for production of different pitches include not only the ability to vary the vocal cord tension when the vocal cords are in the position to produce voiced sounds, but also the ability to produce a continuous, unobstructed air flow. In our next chapter, we shall discuss the question as to whether it is even physiologically possible to vary the tension of the vocal cords in any significant manner when the vocal tract has assumed the configuration of a non-sonorant, voiced or voiceless.

The marking conventions we propose for the pitch features are the following:

TABLE 3.3: THE MARKING CONVENTIONS

I. $$\begin{bmatrix}-\text{son}\\ \alpha\text{voiced}\end{bmatrix} \rightarrow \begin{bmatrix}-\alpha\,\text{highT}\\ \alpha\,\text{lowT}\\ -\text{modify}\end{bmatrix}$$

II. $$\begin{bmatrix}+\text{son}\\ -\text{syl}\end{bmatrix} \rightarrow \begin{bmatrix}-\text{highT}\\ -\text{lowT}\end{bmatrix} \;/\; \#\#\;\underline{\qquad}$$

III. $$\begin{bmatrix}+\text{son}\\ -\text{syl}\end{bmatrix} \rightarrow \begin{bmatrix}\alpha\,\text{highT}\\ \beta\,\text{lowT}\end{bmatrix} \;/\#\#\begin{bmatrix}-\text{syl}\\ \alpha\,\text{highT}\\ \beta\,\text{lowT}\end{bmatrix}\;\underline{\qquad}$$

IV. $$[\underline{u}\ \text{highT}] \rightarrow \left\{\begin{array}{ll}[-\text{highT}] \;/\; \underset{+\text{syl}}{\underline{\qquad}} & \text{(a)}\\ [\alpha\text{highT}] \;/\; \begin{bmatrix}+\text{son}\\ \alpha\,\text{highT}\end{bmatrix}\;\underline{\qquad} & \text{(b)}\end{array}\right.$$

V. $$[+\text{highT}] \rightarrow [-\text{lowT}]$$

VI. $$[\underline{u}\ \text{lowT}] \rightarrow \left\{\begin{array}{ll}[-\text{lowT}] \;/\; \underset{+\text{syl}}{\underline{\qquad}} & \text{(a)}\\ [\alpha\,\text{lowT}] \;/\; \begin{bmatrix}+\text{son}\\ \alpha\text{lowT}\end{bmatrix}\;\underline{\qquad} & \text{(b)}\end{array}\right.$$

VII. $$[+\text{lowT}] \rightarrow [-\text{highT}]$$

VIII. $$\begin{bmatrix}-\text{highT}\\ -\text{lowT}\end{bmatrix} \rightarrow [-\text{modify}]$$

IX. $$[\underline{u}\ \text{modify}] \rightarrow \left\{\begin{array}{ll}[-\text{modify}] \;/\; \underset{+\text{syl}}{\underline{\qquad}} & \text{(a)}\\ [\alpha\text{modify}] \;/\; \begin{bmatrix}+\text{son}\\ \alpha\text{modify}\end{bmatrix}\;\underline{\qquad} & \text{(b)}\end{array}\right.$$

Convention (I) states that voiced non-sonorants are low toned, and voiceless non-sonorants are high toned.

Convention (II) states that initial non-syllabic sonorants, i.e. nasals and glides, cannot be contrastive with respect to pitch. As we have stated before, contrastive pitch has always been noted to be the unique property of the syllabic nucleus. This fact may be paralleled by the not infrequent observation the morpheme initial nasals behave more like obstruents, while post-vocalic nasals are obviously not like obstruents. The contrast between initial and final nasals is apparently sometimes so great that it has led some linguists to conjecture that they are different phonemes. For example, Egerod, in his description of the Lungtu dialect, one of the southern Chinese dialects related to Amoy, finds the obstruent quality of the initial nasals to be so strong that he in fact describes them as prenasalized stops, rather than as nasals, thus claiming that there are no initial nasals in Lungtu.[14] We know that these must be the initial nasals of the other related dialects, and can only suppose from his description that they have either changed diachronically into the homorganic stops, remaining nasalized, or that his fine phonetic ear simply noted the strong contrast between initial and final nasals and that he has chosen to make the distinction clear in this way. Because of the proliferation in the literature of distinctions of this type, and because we have not found any spectrographic evidence to support a claim that the initial nasals or glides form any significant part of the distinctive pitch contour of a syllable, we feel that the pitch on these segments should be specified as mid, or $\begin{bmatrix} -\text{highT} \\ -\text{lowT} \end{bmatrix}$. We have chosen this configuration rather than $[+\text{highT}]$ or $[+\text{lowT}]$ because we do not feel that in the case of initial nasals or glides, that there is a bias either way. It is particularly striking that this is the case when the initial nasal or glide is also phrase initial. When, through morphological or syntactic processes, the lexically morpheme initial nasals or glides become medial in a phonological word, then it is no longer necessarily the case that the pitch of these nasals and glides is $\begin{bmatrix} -\text{highT} \\ -\text{lowT} \end{bmatrix}$. That this is at least sometimes definitely not the case has already been shown in the previous discussion of the Mandarin neutral tone syllables such as /men/ in the noun phrase /ta men/ 他, 們 they (or them).

Convention (III) states that when the nasals or glides are part of the initial consonant cluster, the pitch is assimilated from the preceding consonant. It is, admittedly, a moot question, whether this statement is correct, or whether the assimilation should be from the following syllabic. Our reason for feeling that (III) is the correct statement derives from the fact that in stating the assimilation this way, we can better account for the "dips" and "humps" observed by Kratochvíl in pitch contours, since these "dips" and "humps" are particularly present when there is a pre-vocalic glide. Moreover, we tend to feel that with respect to tone, the prevocalic glide behaves more like a part of the initial consonant cluster rather than as a part of the syllabic nucleus. In the case of post consonantal nasals, it has been the traditional view that the nasal is part of the initial consonant cluster, so that the convention assimilating them tonally to the preceding consonant simply carries out that notion. Since the behavior of prevocalic glides does not differ tonally from that of the nasals in initial consonant clusters, the convention assimilating them in pitch to the preceding consonant would seem, therefore, to be correct. More empirical investigation will, of course, be needed to establish the correctness of this claim unquestioningly.

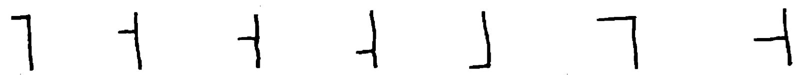
Conventions (V) and (VII) simply state that no segment is specified as $\begin{bmatrix}\text{+highT} \\ \text{+lowT}\end{bmatrix}$.

Convention (VIII) states that there is no segment specified as $\begin{bmatrix} \text{-highT} \\ \text{-lowT} \\ \text{+modify} \end{bmatrix}$. We have already discussed the significance of this restriction at the beginning of this chapter.

Conventions (IVa), (VIa), and (IXa) specify that the unmarked value of the pitch features for syllabic segments is - , or, in other words, that the unmarked tone is the mid tone. Conventions (IVb), (VIb), and (IXb) state that the unmarked value for the sonorant segment following a syllabic is not necessarily the mid tone, but rather the same pitch specification as the syllabic segment. Thus, these conventions state that the long level tones are not to evaluated as being terribly costly, even if the starting point is highly marked. We do not feel that the sonorant segments following the syllabic can be evaluated in a marking system in any manner except relative to the syllabic, the first significant segment of a syllable's pitch contour. We feel that this type of convention ought to be explored to determine if it used for all sonorant features. That is, if one encounters morphemes with VV sequences, it should be the case that when one evaluates the lexicon, the VV sequence consisting of featurewise identical vowels, except in the feature [syllabic] , should be less highly marked than a diphthong, for example. Thus a lexical entry which has the sequence /oo/, for example, should be less marked than one having /oi/ or /oa/. Notice, however, that without these sequential marking conventions, the latter is, in fact, the least marked sequence of the three.[15]

Conventions (IV), (IV), and (IX) are each disjunctively ordered with respect to parts (a) and (b). By virtue of this fact, no [+syl] segment can therefore be affected by part (b), whereas part (b) will affect <u>all</u> [-syl] sonorants following the syllabic. Thus, if the syllabic nucleus consisted of, say, a diphthong and a final nasal, or a geminate vowel and a nasal, one then expects that the least marked contour will be a level contour. We feel that this is a correct consequence of the marking conventions. The effect of these conventions will now be shown for the various pitch contours given in Table 3.1. Recall that these are subsets of matrices of sonorants and sonorant sequences, the first member of which is assumed to be [+syl] and the others [-syl].

TABLE 3.4



high tone	m	m	u	u	u	m u	u u
low tone	u	u	u	m	m	u u	u u
modify	u	m	u	m	u	u u	u u
complexity	1	2	0	2	1	1	0

high tone	u u	u m	u m	m u	u u	u u
low tone	m u	u u	m	u u	m u	m m
modify	u u	u u	u u	m m	m m	m u
complexity	1	1	2	3	3	3

high tone	u m	m m	u u	m u	m
low tone	u u	u u	u m	u u	u m
modify	u m	u u	u u	u m	u u
complexity	2	2	1	2	2

high tone	u u u	u u u
low tone	u m m	m u m
modify	u u u	u u u
complexity	2	2

<u>3.2.2</u> Notice that for the two sequences marked with the superscript (1), we have left one of the spaces blank. This is due to the fact that we wish to emphasize that the specification of the value of this feature in these two cases is an automatic consequence of conventions (V) and (VII), if these two conventions are not ordered with respect to each other or to the other conventions. If the conventions are considered to be ordered, however, and ordered as they have been stated, then the rising tone can still have the specification filled in for [low tone] in the second segment by convention (V), but the specification of [high tone] in the second segment of the falling tone cannot be filled in by convention (VII) since the convention for [u high tone] , convention (IV), precedes convention (VII). One solution is, of course, to order (V) and (VII) before all the other conventions. This would yield the proper specification in all of the cases given. However, we bring the discussion up because there seems to be no motivation, apart from the minor point just stated, for ordering the conventions for [highT] and [lowT] in any way relative to each other.

Convention (IX), on the other hand, seems to be justifiably ordered after the conventions for [highT] and [lowT] , since the specification of this feature is subordinate to that of the others. The point here is simply that the ordering of marking conventions have sometimes been considered as a reflection of the hierarchy of the features. In the case of [highT] and [lowT] , however, there does not seem to be a dominance of either of the features over the other. They seem, rather, to exist on the same level in the hierarchy of pitch features. This equality is reflected by the equivalence of the evaluation measure, which states that a segment marked [+highT] and one which is marked [+lowT] are equally valued. We feel that this is correct and that the complexity measure which states that in a language which has a two way pitch contrast, this contrast is optimally reflected either as one with a high tone ($\begin{bmatrix}\text{+highT}\\\text{-lowT}\end{bmatrix}$) and a non-high or "mid" tone, $\begin{bmatrix}\text{-highT}\\\text{-lowT}\end{bmatrix}$, or as one with a low tone ($\begin{bmatrix}\text{-highT}\\\text{+lowT}\end{bmatrix}$) and a non-low or "mid" tone, $\begin{bmatrix}\text{-highT}\\\text{-lowT}\end{bmatrix}$. As we have already stated in earlier discussions, although languages which have a two way pitch contrast traditionally describe the tones as "high" and "low", one must also remember that pitch is always relative. Hence, the distinctions which are being made are those between high and non-high and low and non-low. The particular contrast must be determined for each language. This type of description, that is, making the distinction one between high and non-high (or mid), rather than high and low, for example, accounts for statements which one sometimes encounters of the type: "The low tone of X language is really the mid tone of Y language". Such a statement has been heard, for example, of the low tone of Burmese and the mid tone of Jingpho Kachin.[16]

Let us now assume that the complexity of a system is equal to the sum of the marked features of its members. Then the simplest system with a three pitch contrast on single segments is the system ˥ ˧ ˩ , or, in terms of features, the system $\begin{bmatrix}+\text{highT}\\-\text{lowT}\end{bmatrix}$, $\begin{bmatrix}-\text{highT}\\-\text{lowT}\end{bmatrix}$, and $\begin{bmatrix}-\text{highT}\\+\text{lowT}\end{bmatrix}$.

For systems which have two-pitch sequences, the systems (a) ˥ ˧ ˩ and (b) ˥ ˧ ˧˥ are considered equally valued by the evaluation criterion. Both types of system exist, but we feel that the system (a) is simpler than (b). There is so far no evidence to show overwhelmingly that this is true, but systems of the type (a) do seem to be more numerous than systems of the type (b). How can we reflect the relative simplicity of (a) as opposed to (b)? We can do this by adding the following statement for the evaluation of systems:

(A). All things being equal, a system which has no sequences containing segments marked for pitch other than the first segment in a sequence is to be preferred to a system in which segments other than the initial segment of a sequence are marked for pitch features.

(Recall that the sonorant sequences referred to here are only those which comprise the syllabic nucleus.)

The evaluation criterion, with or without statement (A), which is irrelevant here, states that systems such as (c) ˥ ˧ ˧˥ and (d) ˧ ˩ ˧˩ are equally valued. For systems which have only two contrastive pitch heights, and which have three distinctive sequences, including one contour sequence, we see no reason to favor (c) over (d) or vice-versa. When there is a system with three two-pitch sequences, where one of the sequences is a contour, one finds that both (c) and (d) occur. The fact that they do not often occur, we feel, is reflected by statement (A).

Statement (A), however, is clearly not sufficient to adjust the evaluation criterion. Consider the following sets of five different two pitch sequences: (e) ˥ ˥˧ ˧ ˨ ˩ (f) ˥ ˧ ˩ ˩˧˩ ˥˧˥ and (g) ˥ ˧ ˩ ˧˥ ˥˧, where the last two members of the set (g) are specified by the feature matrices, $\begin{bmatrix} \text{-highT} \\ \text{-lowT} \\ \text{-modify} \end{bmatrix} \begin{bmatrix} \text{+highT} \\ \text{-lowT} \\ \text{+modify} \end{bmatrix}$, and $\begin{bmatrix} \text{-highT} \\ \text{-lowT} \\ \text{-modify} \end{bmatrix} \begin{bmatrix} \text{-highT} \\ \text{+lowT} \\ \text{+modify} \end{bmatrix}$, respectively. Clearly we wish to say that of the three systems, system (f) is preferable to both systems (e) and (g), and that (e) is preferable to (g). The evaluation measure, without statement (A), would state that these three systems are equally valued. Even with statement (A), we get the wrong result that (e) is preferred to (f) and (g). Hence, we add the following two statements:

(B). No sequence can be marked for the feature [modify] unless some other sequences in the set are marked for [highT] only and for [lowT] only.

(C). All other things being equal, a system in which more features have only the specification u is preferable to a system in which fewer features have only the specification u.

Statement (B) emphasizes the order of the hierarchy between the equally ranked features [highT] and [lowT] and the feature [modify]. We feel that this reflects the fact that there are no languages known which make a lexical distinction between a high tone and a high-mid tone which does not also have a low tone. Statements (B) and (C) together determine that system (f) is preferred to the otherwise equally valued systems (e) and (g). Statement (A) would determine that (e) is preferred to (g). This last consequence is probably the strongest evidence to date that statement (A) is needed and is correct.

One peculiarity of the marking conventions is that the sequences in Table 3.4 marked with the superscript (2) are evaluated as being more costly than those with a superscript (3). This is due to the fact that the conventions as they stand favor sequences which begin at the "mid" point, $\begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix}$, rather than those which begin from a marked position. There is, as one can readily see, no other difference in these contours. Whether this favoring is correct or not is a completely open question. The answer can only be determined by a detailed examination of those languages in which there are five distinctive pitch heights and which also have pitch sequences in single syllables. Because such material is not now available, we can make no decision and will leave these evaluations as they stand. If it should be the case, after further investigation, that these evaluations are incorrect, we must point out that this does not imply that the entire set of conventions is incorrect, but only that a more complicated set of conventions is involved when the feature [modify] is specified as m.

3.2.3 The original motivation behind the concept of marking conventions was not only the desire to be able to specify "optimal" sets of segments, but also to reflect optimal types of segment changes in the evaluation of rules. This latter role of the marking conventions has been termed "linking".[17] The notion which had not been hitherto captured by the simple feature count method of rule evaluation had been the difference in "naturalness" of a rule such as

(15). i → u as opposed to (16). i → ɨ . The feature count method of evaluation gave the result that (16) was less expensive than (15), while (15) seemed to be the more "natural", or statistically common, rule. To overcome this, the idea of using marking conventions as rules which apply throughout the phonology was introduced. We wish to do the same with the conventions which we have just introduced for the pitch features. The problem, however, lies in the fact that up to this moment, phonological rules involved changes in the feature specification of one segment per rule only, with the exception of rules for metathesis. If we hold by this restriction on the form of phonological rules, then how can part (b) of conventions (IV), (VI), and (IX) be used as linking rules? To clarify the question, let us consider a hypothetical situation in which a long level high tone changes into a long level mid tone. Suppose also that the sequences are two-pitch sequences. A graphic representation of this alternation is: ˥ → ˧ . If we were to restrict ourselves to rules changing the feature values of only one segment at a time, the alternation can be expressed by the following rules.

(17). [+highT] → [-highT] / ____ [+highT]

(18). [+highT] → [-highT] / [-highT] ____

But let us consider another situation in which the long high level tone changes into a mid to low falling tone, graphically represented as ˥ → ˧˩ . This situation could be expressed by the following rules.

(19). [+highT] → [-highT] / ____ [+highT]

(20). [+highT] → [+lowT] / [-highT] ____

The two sets of rules are evaluated as equally costly, for we have not been able to make use of convention (IVb) as a linking rule. However, the situation described by rules (17) and (18), we feel, is one that ought to be evaluated as less costly than the one described by rules (19) and (20).

Let us digress momentarily and consider a parallel situation in segmental phonology with languages which represents all phonetic long vowels as geminate clusters. Suppose that such a language has an alternation between /oo/ and /uu/ under certain conditions. We then encounter the same problem of how to formally state a rule or rules to express this alternation. If phonological rules are restricted to changing the feature specification of only one segment at a time, then the alternation between /oo/ and /uu/ might be expressed as the following:

(21). o → u / ___ o

(22). o → u / u ___

However, let us consider another possible alternation, that between /oo/ and /ua/. This latter alternation might be formulated as:

(23). o → u / ___ o

(24). o → a / u ___

In all of these rules, (21) - (24), if one uses the marking conventions for vowels in SPE, there is only one feature involved on the right side of the arrow. Thus, the evaluation criterion does not differentiate between the two alternatives in terms of "naturalness". In particular, even using a theory of markedness which considers only isolated segments, there is no way of stating that the alternation expressed by

(21) and (22) is somehow more natural than that of (23) and (24). On the other hand, if the theory of marking for segmental features other than the prosodic features were extended to include sequential conventions, then it might be possible to use these extended notions to capture the difference between the alternation expressed by (21) and (22) as opposed to that expressed by (23) and (24).

We suggest that one solution to both the tone and vowel problem would be to make use of the formalism developed to handle cases of metathesis in situations like these. Then the tone changes expressed by (17) and (18) could be collapsed to the following:

(25).
$$\overset{1}{\begin{bmatrix}+\text{son}\\+\text{highT}\end{bmatrix}} \quad \overset{2}{\begin{bmatrix}+\text{son}\\+\text{highT}\end{bmatrix}} \rightarrow \overset{1}{\begin{bmatrix}-\text{highT}\end{bmatrix}} \quad \overset{2}{}$$

In (25), no features need be specified for the second segment on the right side of the arrow, for the interpretation of the rule will be that the marking conventions will apply for the features which have been mentioned, namely, [highT] . Part (b) of convention (IV) will apply and specify the value of [highT] in the second segment as [-highT] . The evaluation measure then can remain a simple feature count. If we introduce this use of bisegmental rules, then we can readily reflect the hierarchy of "naturalness" in tone sandhi.

Sandhi which only affects one segment, such as the change ˥ → ˦ will be the least expensive, and can be formalized as:

$$(26).\quad [+highT] \rightarrow [-highT] / \_\_\_\_ [+highT]$$

Sandhi of the type ˥ → ˧ is more expensive (<u>cf</u>. rule (25)), and changes such as ˥ → ˨˩˦ will be the most expensive (<u>cf</u>. rules (19) and (20)). One might question whether ˥ → ˨˩˦ should be expressed as a two rule change, (19) and (20), or whether it should be expressed as:

$$(27).\quad \overset{1}{\begin{bmatrix} +son \\ +highT \end{bmatrix}} \overset{2}{\begin{bmatrix} +son \\ +highT \end{bmatrix}} \rightarrow \overset{1}{[-highT]} \overset{2}{[+lowT]}$$

thus making use of the same formalism as in (25). We have no strong evidence to make the decision either way, and so leave it as an open question. One may wish to consider as partial support for the formulation (27) as opposed to the set (19) and (20), the fact that in order to capitalize on the linking role of the conventions, we must make use of the bisegmental rule formalism to express the change ˨˩˦ → ˧ . This latter can be formalized as:

$$(28).\quad \overset{1}{\begin{bmatrix} +son \\ +highT \end{bmatrix}} \overset{2}{\begin{bmatrix} +son \\ +lowT \end{bmatrix}} \qquad \overset{1}{[-highT]} \quad \overset{2}{}$$

with the same interpretation of the right side as in (25). Not only must this be formalized as (25) is, but we feel that it should be evaluated as less expensive than sandhi of the type ˥ → ˧˩ . Whether this is to be reflected by the fact that the latter change must be formalized as a two part rule (19) and (20), or by the fact that in the statement of (27), the feature specification of the second segment on the right side must be directly stated rather than automatically specified by the marking conventions, we cannot now say.

Notice that this hierarchy of tone changes, (26), (25), (28), and (27) has its direct correlate in rules affecting vowels. That is, we feel that the same hierarchy is true of vowels, /aa/ → /ai/ being more natural than /aa/ → /ææ/, which is in turn either more natural or equally as natural as /au/ → /oo/, which is more natural than /aa/ → /ei/. The tonal hierarchy is not strictly held between (25) and (28), that is, the evaluation of such rules is equivalent, and at this point, we see no reason for a ranking between them.

The marking conventions presented are, of course, highly tentative. We feel that they are correct for languages which have no lexically distinctive long vowels. However, it is quite probable that the sequential conventions are more complex than those stated. This can only be determined by a more detailed study of many more languages.

3.3 Let us now consider three Chinese dialects, Amoy, Chaochow, and Lungtu, to determine what some of the more complicated tone systems are like, as well as what kind of sandhi exists and what are some of the situations in which sandhi takes place.

The tone system of Amoy is traditionally described as consisting of five "long" tones and two "short" tones, where a "long" tone is a two pitch sequence occurring on syllables with a long vowel, a diphthong, or a vowel plus nasal as its nucleus and not closed by a stop. The "short" tones are single, or unit, pitch sequences found on syllables closed by stop consonants since these have short syllabic nuclei. The long tones are described as high level, mid level, low level, mid to high rising, and high to low falling. The tones of the short syllables are described as high and mid. Bodman describes the tones of the short syllables as high and low,[18] but since the descriptions in both Lo Ch'ang P'ei's Phonetics and Phonology of the Amoy Dialect[19] and the Hanyu Fangyan Cihui[20] give these tones as high and mid, we shall accept the description given in the latter two sources, rather than Bodman's. Unfortunately, Bodman did not use the "tone letters" introduced by Chao, so that it may simply be the case that having found only a two way opposition in the closed syllables, he chose to call the pitch levels "high" and "low" although recognizing that the "low" was just another name for [-highT].

These tones are normally called the "isolation" tones, or the forms of the pitch contour when the syllable is in phrase final position. In addition to this set, there is another set, called the "combination" tones, which are the pitch contours of the syllables when in non- phrase final position. Traditional grammars usually give a list of rules specifying the correlation between the "isolation" tones and the "combination" tones. Such a list is given by Bodman as the following:[21]

> "A form with isolation high tone, if not ending in /p/, /t/, /k/, or /q/, has mid tone in combination.
>
> A form with isolation rising tone has mid tone in combination.
>
> A form with isolation mid tone has low tone in combination.
>
> A form with isolation low tone, except for /p/, /t/, /k/ finals has falling tone in combination.
>
> A form with isolation falling tone, has high tone in combination.
>
> A form with isolation high tone having stop final (p, t, k, or q), has low tone in combination.
>
> A form with isolation low tone with /p/, /t/, /k/ finals has high tone in combination."

Note: Bodman uses /q/ as a symbol for a lax glottal stop. Recall, too, that the "low" mentioned in the last two statements is what we are calling a mid tone.

The conditions under which these principles of tone change apply are the following:[22]

> "As far as tone change is concerned, the principles hold whether a form appears in a free construction i.e. <u>chà</u> 'to cook in oil' in <u>châ cǐt puā: mī</u> 'fry a dish of noodles', or whether it appears in a more limited construction of compound words, such as the same form in <u>châ-mī</u> 'fried noodles'. In both cases, the original falling tone of the form spoken in isolation: <u>chà</u>, is replaced in combination by the form with the high tone: <u>châ</u>."

Note: The tone marks used here represent the following: ` = falling tone, ^ = high tone, ˇ = low tone, and ¯ = mid tone. The mark /:/ is used to indicate a nasalized vowel, not a long one.

In order to make the discussion more clear, the feature matrices of the isolation tones are given in Table 3.5 below and the matrices of the combination forms are given in Table 3.6 below. The numbering of the contours in each table indicates the correspondence between the various matrices.

TABLE 3.5: AMOY ISOLATION TONES

	(1) ˥˥	(2) ˧˥	(3) ˧˧	(4) ˩˩	(5) ˥˩	(6) ˥	(7) ˧
highT	+ +	- +	- -	- -	+ -	+	-
lowT	- -	- -	- -	+ +	- +	-	-
modify	- -	- -	- -	- -	- -	-	-

TABLE 3.6: AMOY COMBINATION TONES

	(1,2) ˧˧	(3) ˩˩	(4) ˥˩	(5) ˥˥	(6) ˧	(7) ˥
highT	- -	- -	+ -	+ +	-	+
lowT	- -	+ +	- +	- -	-	-
modify	- -	- -	- -	- -	-	-

In Table 3.6, the first sequence of matrices is marked (1,2) since both the isolation high tone syllables and the isolation rising tone syllables have a mid tone in combination.

Combination tones occur whenever the syllable is not phrase final, irrespective of the structure of the phrase, and of the syntactic relations of the syllable, as expressed by bracketing, to the other syllables in the phrase. There is one exception to the preceding statement and it is the following. If a syllable has been assigned [-stress] , then irrespective of whether that syllable is phrase final or not, the preceding syllable is articulated with its isolation pitch contour, not with its combination pitch contour. We do not wish to discuss the stress rules, and therefore will simply state that among items which are unstressed are particles and directional complements to verbs. Thus, if in a phonological phrase, there is a verb phrase consisting of a monosyllabic verb followed by some directional adverbial, the adverb will be unstressed and the verb will always occur with its isolation pitch contour, irrespective of the fact that it is not phrase final.

Let us now discuss the rules which express the alternation between the contours in Table 3.5 and those in Table 3.6 under the conditions just stated. One question which we must consider is, which form of the pitch contour, "isolation" or "combination", is to be specified in the lexicon for a given syllable? The traditional view has been to consider the "isolation" tone basic, and to derive the combination tones from them. This view is apparently also held by Wang in his article, "Phonological Features of Tone", where he specifies the changes stated above, except for (6) and (7), the closed syllable changes, as a single phonological rule involving the variables α and β . Before stating his rule, we must first explain that Wang represents all occurrences of the low level tone, ˩ , as a mid to low falling tone, ˧˩ . We will discuss the motivation for this representation later. Keeping in mind that the representation of (4) in Table 3.5 and (3) in Table 3.6 is ˧˩ , specified by Wang's features as $\begin{bmatrix} +\text{high} \\ +\text{fall} \\ -\text{rise} \\ +\text{Long} \end{bmatrix}$, the rule he states to account for the alternation from isolation to combination forms is:[23]

$$(29).\quad \begin{bmatrix} \alpha\,\text{high} \\ \beta\,\text{fall} \end{bmatrix} \rightarrow \begin{bmatrix} \beta\,\text{high} \\ -\alpha\,\text{fall} \end{bmatrix}$$

The environment of (29) must be stated as:

$$(30).\quad \underset{\begin{array}{l}+\text{stress}\\ +\text{Long}\end{array}}{\underline{\qquad\qquad}} \quad \left(C\begin{bmatrix} V \\ +\text{stress} \end{bmatrix} V\right)_1 \begin{Bmatrix} \big]_{\text{phon. phrase}} \\ C\begin{bmatrix} V \\ -\text{stress} \end{bmatrix} \end{Bmatrix}$$

Here $\left(C\begin{bmatrix} V \\ +\text{stress} \end{bmatrix} V\right)_1$ is an abbreviation for one or more stressed syllables. The feature [Long] in (30) does not refer to vowel length, but rather is a diacritic used by Wang to differentiate between syllables closed by a stop consonant and those not closed by a stop. [+Long] , therefore, refers to the open syllables, which category includes those with a nasal as the final segment. [-Long] is used to specify those syllables closed by a stop. The sandhi rule, (29), applies to syllables which are stressed and not closed by an obstruent, whenever these syllables occur before one or more stressed syllables within a phonological phrase or before an unstressed syllable. For example, consider a string of syllables, S_1 S_2 S_3 S_4 S_5 S_6 , where each S_1 is a [+Long] syllable. Suppose also that S_4 is unstressed and that the end of the phonological phrase occurs after S_6. Rule (29) will then apply simultaneously to S_1, S_2, and S_5 , as determined by the environment (30).

In order to explain how rule (29) works, we will now give the feature matrices of the Amoy open syllables as presented in Wang's paper.[24]

TABLE 3.7: AMOY TONES (WANG)

	(1) ˥	(2) ˧˥	(3) ˧	(4) ˨˩	(5) ˥˧
high	+	+	-	-	+
fall	-	-	-	+	+
rise	-	+	-	-	-
Long	+	+	+	+	+

The numbers in Table 3.7 correspond to those in Table 3.5.

Rule (29), then states that the value of [high] of the derived tone is that of [fall] of the basic tone, and that the value of [fall] of the derived tone is opposite to the value of [high] of the basic tone.

A single rule must express these alternations because the changes can be seen to proceed in a circular fashion, that is, (1,2) $\longrightarrow$ (3) $\longrightarrow$ (4) $\longrightarrow$ (5) $\longrightarrow$ (1). It is from this fact that we find the reason for Wang's abstract representation of the low level tone as (4) in Table 3.7. In his system of features, there is only one major opposition of pitch heights and that is the one between high and non-high. Thus, a tone specified with the matrix $\begin{bmatrix}\text{-high}\\ \text{-contour}\end{bmatrix}$ may abstractly represent either the mid tone, as seen in Amoy, or the low tone, as shown in his table of tones and their features.[25] However, this matrix cannot represent both configurations simultaneously. Were he to represent the tones of Amoy as the configurations we give in Table 3.5, he would have to specify the mid level tone as $\begin{bmatrix}\text{-high}\\ \text{-fall}\\ \text{-rise}\\ \text{+mid}\\ \text{+central}\end{bmatrix}$ in order to keep it distinct from the low level tone, which is specified as $\begin{bmatrix}\text{-high}\\ \text{-fall}\\ \text{-rise}\\ \text{-mid}\\ \text{-central}\end{bmatrix}$.

However, if the mid tone were specified in the manner just stated, then clearly rule (29) could not be used to express the sandhi. A much more complex rule would have to be stated, primarily because the relations of the features [mid] and [central] to the other features in Wang's system are quite complex. Thus, he chooses to represent the low level tone as a low falling tone, with the result that (29) will correctly express the alternations. According to all the sources, however, the low level tone is a low level tone, and not one which is falling.[26] Hence, if one accepts Wang's analysis and rule (29), one will have to add the following two rules to the phonology to obtain the correct phonetic contours of the Amoy tones.

(31). $\begin{bmatrix} \text{-high} \\ \text{-rise} \\ \text{-fall} \end{bmatrix} \rightarrow \begin{bmatrix} \text{+mid} \\ \text{+central} \end{bmatrix}$

(32). $\begin{bmatrix} \text{-high} \\ \text{+fall} \end{bmatrix} \rightarrow \begin{bmatrix} \text{-fall} \end{bmatrix}$

Rule (31) states that all non-high non-contour tones are also $\begin{bmatrix} \text{+mid} \\ \text{+central} \end{bmatrix}$, that is, the non-high non-contour tone is now specifically identified as a mid level tone. Rule (32), which must follow (31) for obvious reasons, then states that the non-high falling tone is a non-high level tone, the normal feature specification of a low level tone in Wang's system.

Let us now consider the tones of Amoy to be specified as in Table 3.5 and Table 3.6. There we see that the difference between "long" and "short" syllables never need be explicitly stated, for the "long" syllables

are precisely those which have the structure $C_o(G)\ V \begin{Bmatrix} V \\ N \end{Bmatrix}$,
while the "short" syllables are those with the structure $C_o(G)\ V\ C$. Thus, one could say that the difference between the two syllable types is expressed by the fact that the "long" syllables have a complex syllabic nucleus, that is, one which consists of either a long vowel (a geminate vowel), a diphthong, or a vowel plus nasal cluster. The "short" syllables, on the other hand, are characterized by the fact that the syllabic vowel is always followed by a stop. Keeping this in mind, let us now state the rule for Amoy sandhi when the lexical tones are considered to be those in Table 3.5, and the derived tones to be those in Table 3.6.

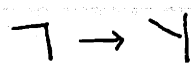
$$(33)\quad \begin{bmatrix} +\text{son} \\ \alpha\text{lowT} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\text{highT} \\ \langle -\alpha\text{lowT} \rangle \end{bmatrix} \Big/ \left\{ \begin{array}{ll} \begin{bmatrix} -\text{lowT} \\ +\text{stress} \end{bmatrix} \langle \alpha\text{highT} \rangle \underline{\qquad\qquad} & \text{(a)} \\ \underline{\qquad\qquad} \\ +\text{stress} \begin{bmatrix} +\text{son} \\ -\text{highT} \\ \langle \alpha\text{lowT} \rangle \end{bmatrix} & \text{(b)} \end{array} \right.$$

The rest of the environment for (33) is the same as for Wang's rule, namely

$$/ (C \begin{bmatrix} V \\ +\text{stress} \end{bmatrix} [+\text{seg}])_1 \left\{ \begin{array}{l}]_{\text{phon. phrase}} \\ C \begin{bmatrix} V \\ -\text{stress} \end{bmatrix} \end{array} \right.$$

In applying (33), we assume as before, that only the syllabic vowel can be specified as [+stress]. Thus it follows that (33) can only apply to the "long" syllables, because only these syllables satisfy the requirement that there is a sonorant immediately following the stressed vowel. Notice that the feature [stress] must always be mentioned irrespective of which rules one uses, for only the stressed syllables are ever affected. Thus, using our feature representation of the tones, the feature [Long] need never appear, for (33) is automatically restricted to applying to the long syllables by virtue of the way in which the environment is stated.

Rule (33a) applies to the pitch sequences which begin with a [-lowT] segment. Hence, it applies to contours (1), (2), (3), and (5) on Table 3.5, the high level tone, the mid to high rising tone, the mid tone, and the high to low falling tone. (33a) states that the second segment in such sequences takes on the value for [highT] that it had for [lowT] in the isolation forms. If the second segment happens also to be specified as $\begin{bmatrix} \alpha\,\text{highT} \\ \alpha\,\text{lowT} \end{bmatrix}$, or $\begin{bmatrix} -\text{highT} \\ -\text{lowT} \end{bmatrix}$, in other words, then the angle bracket notation states that the derived matrix of the second segment has the opposite value for the feature [lowT] as the matrix for the citation form. Applying (33a) to the contours in Table 3.5, we get the following derivations:

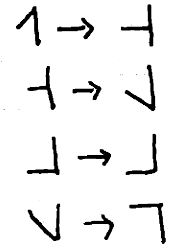
(34). 

(35).

(36).

(37). ((33a) inapplicable)

(38).



(33b), which is conjunctively ordered with respect to (33a), then applies to the output of (33a). It will apply only to the sequences in which the second segment is [-highT] . Thus, it will not apply to the derived form in (38). The rule states that the derived matrix has the same specification for [highT] as the input matrix had for [lowT] . Moreover, the angle brackets state that if the following segment is [+lowT], then the derived matrix takes on the opposite value for [lowT] as the input matrix had. Applying (33b) to the derived contours (34) - (38), we get the following derivations:

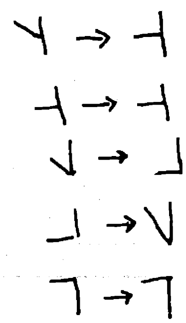
(34a).

(35a).

(36a).

(37a).

(38a). ((33b) does not apply to ).



Thus, if the citation forms are considered to be the lexical forms, it is clear that (33) will apply to the contours given in Table 3.5 to give the correct derived contours of Table 3.6. (33) is admittedly more cumbersome than (29). However, one must recall that in order to state (29), a more abstract representation of the low level tone is required, as well as the use of the feature [Long] . In addition, one must state rules (31) and (32) to obtain the correct phonetic contours. If the Amoy tones are represented as we suggest, then none of these complications arise, for then one rule, (33), can express the alternations, and no abstract forms or adjustment rules are needed. Moreover, one need not use the diacritic [Long] , for the necessity for this diacritic arises solely from the fact that Wang considers the tone features to be syllable features and therefore has no way of determining if a syllable is closed by an obstruent or not. Using the segmental analysis of tone, one can easily distinguish between syllables closed by obstruents and those not.

Let us now return to the question of which pitch specification is given in the lexicon for a given syllable in Amoy, the "isolation" form, or the "combination" form? The solution just presented assumes that the isolation forms, those given in Table 3.5, are the lexical forms and that the sandhi rule derives the contours in Table 3.6. However, it is not obvious that the isolation contours must be considered as lexical, for recall the problems which arise in Mandarin phonology if the citation form of the third tone is taken to be lexical. Thus, let us now consider how the sandhi might be expressed if the combination contours are considered lexical.

Clearly, since there are only four contrastive two-pitch sequences in non-final position, and five in phrase final position, the lexical tone contours cannot be just the four sequences given in Table 3.6. That is, recall that both the syllables with high level tone contours and those with rising tone contours in phrase final position have the same pitch contour in non-phrase final position, a mid level contour. Thus, in order to be able to consider the isolation contours as the derived contours, we must give a more abstract lexical representation to either the phonetic mid tone which alternates with the phrase final high tone, or to the phonetic mid tone which alternates with the phrase final rising tone. Suppose we represent the phonetic mid tone which alternates with the phrase final high tone as a high to mid falling tone. Thus, rather than just the four contours given in Table 3.6, we have the five contours given below:

TABLE 3.8: AMOY NON-FINAL LEXICAL CONTOURS

	(1) ˥˧	(2) ˧˧	(3) ˩˩	(4) ˥˩	(5) ˥˥
highT	+ −	− −	− −	+ −	+ +
lowT	− −	− −	+ +	− +	− −
modify	− −	− −	− −	− −	− −

We have left out the "short" contours, since all of the discussion is only concerned with the five long contours. The numbers indicate the contours in Table 3.5 with which these contours alternate in phrase final position. If we consider these contours to be the lexical contours, then the following rule will give the corresponding derived contours, those in Table 3.5, whenever the syllables occur either in phrase final position or before an unstressed syllable.

(39).
$$\begin{bmatrix}+\text{son}\\ \alpha\text{highT}\end{bmatrix} \rightarrow \begin{bmatrix}\alpha\text{lowT}\\ \langle -\alpha\ \text{highT}\rangle\end{bmatrix} / \left\{ \begin{array}{ll} \underset{+\text{stress}}{\underline{\qquad}} \begin{bmatrix}+\text{son}\\ +\text{lowT}\end{bmatrix} & \text{(a)} \\ \begin{bmatrix}+\text{stress}\\ -\text{lowT}\end{bmatrix} \underset{\langle\alpha\text{lowT}\rangle}{\underline{\qquad}} & \text{(b)} \end{array} \right.$$

The remaining environment of (39) is
$$/ \left\{ \begin{array}{l}]_{\text{phon. phrase}} \\ C \begin{bmatrix}V\\ -\text{stress}\end{bmatrix} \end{array} \right.$$

Thus, it is clear that in (39), the environment is simpler than in (29) or (33), for the rule simply applies to the sonorants of syllabes which either end phonological phrases, or are immediately before an unstressed syllable. That is, rather than having to state the remaining environment as

$$/ \left(C \begin{bmatrix} V \\ -\text{stress} \end{bmatrix} [+\text{seg}] \right)_1 \begin{Bmatrix} \quad]_{\text{phon. phrase}} \\ C \begin{bmatrix} V \\ -\text{stress} \end{bmatrix} \end{Bmatrix}$$

we can now simply state it as

$$/ \begin{Bmatrix} \quad]_{\text{phon. phrase}} \\ C \begin{bmatrix} V \\ -\text{stress} \end{bmatrix} \end{Bmatrix} .$$

Rule (39a) can only apply to contours (3) and (4) on Table 3.8, and will give the following:

(40). ˩ → ˧˩

(41). ˥˩ → ˩

(39b) is conjunctively ordered with respect to (39a) and will therefore affect the derived form in (40) as well as contours (1), (2), and (5) on Table 3.8. It does not apply to the derived contour in (41). The derivations are the following:

(42). ˧˩ (from (40) above) → ˧

(43). ˧˨ → ˥

(44). ˧ → ˧˥

(45). ˥ → ˥˩

Looking at the derived forms in each of the terminal derivations, (41) - (45), one sees that the derived contours are exactly those in Table 3.5, as desired.

Rule (39), therefore, expresses the sandhi when the lexical tones are taken to be the non-final contours. In order to get the proper phonetic shape of contour (1) on Table 3.8, we must add one minor rule to the phonology which is:

$$(46).\quad [+\text{son}] \rightarrow \begin{bmatrix}-\text{highT}\\ -\text{lowT}\end{bmatrix} \;/\; \_\_\_ \begin{bmatrix}-\text{highT}\\ -\text{lowT}\end{bmatrix}$$

Rule (46) simply changes contour (1) to one which is identical to contour (2), the mid level tone, whenever it occurs.

We see from the preceding that, from the point of view of formulating rules to express the sandhi, it is not obvious which contours are to be considered lexical, the non-phrase final contours, or the phrase final contours. That is, one can express the sandhi in either direction, from isolation forms to combination forms, or from combination forms to isolation forms. Rule (39), which expresses the sandhi in the latter direction, is somewhat simpler than (33). In particular, the statement of the environment is simpler. Apart from considerations of simplicity in rules, there is another reason to regard the solution using rule (39) as correct, and this arises from a consideration of the Chaochow dialect and its tone sandhi.

3.4 Li Yong-ming, in his Phonology of Chaochow,[27] describes the isolation tones of Chaochow as the following: long mid level (˧, 33), high to mid falling (˥˧, 53), falling-rising (˨˩˧, 213), long high level (˥, 55), mid to high rising (˧˥, 35), long low level (˩, 11), short high (˥, 5), and a short mid (or low?) (˨, 2). We shall call this latter short tone, mid, or $\begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix}$, although it is difficult to determine from the description exactly what it is. Li then procedes to describe the sandhi, since the above configurations are only the isolation contours. He explains that, as in Amoy, whenever a syllable is not phrase final, irrespective of its syntactic relations to the other syllables in the phrase, its tone is not the same, for the most part, as the isolation tone. The changes are the following:

a. The mid tone remains a level mid tone everywhere.

b. The high to mid falling tone changes to a mid to high rising tone before all syllables.

c. The falling-rising tone is changed to a high to mid falling tone before all syllables.

d. The long high tone changes to a low to mid rising tone before all syllables.

e. The mid to high rising tone becomes a mid to low falling tone before all syllables.

f. The long low level tone remains low level everywhere.

g. The closed high tone syllable becomes a mid tone closed syllable in non-final position.

h. The mid tone closed syllable becomes high tone in non-final position.

There is one further change of a lowering of the entire quality of the falling tone of (c) when the syllable precedes syllables not beginning on a high pitch. This, however, is probably a natural consequence of the sequential articulation of these syllables, so that preceding syllables which have a tone sequence beginning with a high pitch, the pitch of the first contour is raised. Li Yong-ming's description of this difference is that the fall is from mid to low rather than from high to mid when the syllable precedes one that begins with a non-high pitch. However, the <u>Hanyu Fangyan Cihui</u>, describes the tone difference more in the degree of the fall, where the fall ends lower before non-high starting contours, and ends higher (at the mid pitch) before syllables with a pitch contour beginning with a high pitch.[28] Because of the discrepancy in the descriptions, and because of the nature of the environment in which the higher allotone is found, we suspect that this may be a natural consequence of the environment of the syllable, rather than a change which must be specified in the phonology. That is, we feel that the raising of the "fall" is probably a phonetic consequence of universal nature because of the environment of the syllable, and hence, like the extra length of English vowels in syllables closed by voiced stops, as opposed to those closed by voiceless stops, needs not to be explicitly stated in the phonology.

For the Chaochow phonology as for the Amoy phonology, we see that the nature of the sandhi rules depends on which pitch contours are specified in the lexicon, the "isolation" forms, or the "combination" forms. In order to help make the following discussion clear, the feature matrices of the isolation tones are given in Table 3.9 below, and the matrices of the combination forms are given in Table 3.10. The numbering of the contours in each table indicates the correspondence between the various matrices.

TABLE 3.9: CHAOCHOW ISOLATION TONES

	(1) ˧	(2) ˥˧	(3) ˨˩˧	(4) ˥	(5) ˧˥	(6) ˩	(7) ˥	(8) ˧
highT	- -	+ -	---	+ +	- +	- -	+	-
lowT	- -	- -	-+-	- -	- -	+ +	-	-
modify	- -	- -	---	- -	- -	- -	-	-

TABLE 3.10: CHAOCHOW COMBINATION TONES

	(1) ˧˧	(2) ˧˥	(3) ˥˧	(4) ˩˧	(5) ˧˩	(6) ˩˩	(7) ˧	(8) ˥
highT	- -	- +	+ -	- -	- -	- -	-	+
lowT	- -	- -	- -	+ -	- +	+ +	-	-
modify	- -	- -	- -	- -	- -	- -	-	-

Let us now determine the consequences of considering the contours in Table 3.9, the isolation contours, to be lexical. There are several objects to considering the isolation contours as the lexical contours, most of which arise from the fact that then one would have to consider contour (3) on Table 3.9 as lexical. The marking conventions do not evaluate this configuration as very costly, the complexity being <u>2</u>, as given in Table 3.4. However, it is a high marked configuration since it can occur only if there is a sequence of three sonorants in the syllable, the first being the syllabic vowel. That is, a sequence of segments $C\begin{bmatrix}V\\+syl\end{bmatrix}\begin{bmatrix}+son\\-syl\end{bmatrix}\begin{bmatrix}+son\\-syl\end{bmatrix}$ is itself a highly marked sequence. Thus, one would prefer not to have to list such sequences in the lexicon. Suppose that we were to say that such sequences did exist in the lexicon for Chaochow. We would then have to complicate both the statement of the canonical form of the syllable as well as the rules in the phonology. That is, the extra-long form can only occur in two positions in a phonological phrase; the first being when it is stressed and occurs before an unstressed syllable, the second being when it is stressed and is the last syllable in a phonological phrase. In all other positions, the syllable is not extra-long. Thus, except for these two occurrences, the Chaochow syllable may be said to have the same canonical form as the Amoy syllable, $C_o(G)\ V \begin{Bmatrix} V \\ N \\ C \end{Bmatrix}$.

If we were to consider the combination contours as lexical, the preceding could be given as the canonical form of all the Chaochow syllables. If the isolation contours are considered to be lexical, then the statement of the canonical form of the syllable must be far more complex, for it would then have to be expanded with certain constraints to accomodate the extra-long forms. Since the Chaochow extra-long syllables are similar to the Mandarin extra-long syllables in that they do not permit the vowel cluster before a nasal to be a diphthong, it is clear that the complications which will arise when the canonical form of the Chaochow syllable is extended to include the extra-long syllables will be similar to those already discussed with respect to the Mandarin syllable structure and the Mandarin third tone in chapter II, section 2.1.5. We will therefore not discuss this problem any further.

More important than the problems connected with the statement of the canonical form of the syllable, however, are the problems which arise in the phonology if the contours in Table 3.9 are considered to be lexical. Let us assume that the shape of the extra-long syllables in the lexicon is $C_0(G)\ V_1\ V_2 \begin{Bmatrix} V \\ N \end{Bmatrix}$, where $V_1 = V_2$ when V_2 is followed by a nasal. Since the combination form of these syllables is not extra-long, it is clear that in addition to rules which apply to syllables with the contours given in Table 3.9 to derive those given in Table 3.10, whenever the syllables are stressed and are not phrase final or before an unstressed syllable, there will also have to be a rule which deletes one of the sonorants of the extra-long syllables under these same conditions. Let us now consider the nature of the deletion rule. The problem is, which sonorant is to be deleted? It cannot be the last sonorant, for the last sonorant may be a nasal, and the nasal appears in both the normal and extra-long forms of the syllable. Suppose it is the second sonorant which is deleted. Recall that the pitch specification of the three sonorants in extra-long syllables was given in Table 3.9 as $\begin{bmatrix} \text{-highT} \\ \text{-lowT} \end{bmatrix} \begin{bmatrix} \text{-highT} \\ \text{+lowT} \end{bmatrix} \begin{bmatrix} \text{-highT} \\ \text{-lowT} \end{bmatrix}$. If we were to simply delete the second sonorant, then the pitch configuration would be $\begin{bmatrix}\text{-highT}\\\text{-lowT}\end{bmatrix}\begin{bmatrix}\text{-highT}\\\text{-lowT}\end{bmatrix}$, identical to contour (1) in Table 3.9, the mid level tone. One could not, then, derive the combination form of these extra-long syllables since they would be non-distinctive in their pitch specification from syllables with underlying mid level tones. Recall that the mid level tone remains a mid level tone in isolation and in combination, so that there is no way of ordering rules to avoid this problem. Thus, if one were to say that the sonorant to be deleted is the second vowel, one must first formulate a rule which will change the pitch of the first sonorant, the syllabic, so that the contour after the second sonorant is deleted remains distinctive.

Alternatively, one could say that the syllabic vowel is deleted. The resulting contour would be $\begin{bmatrix}\text{-highT}\\\text{+lowT}\end{bmatrix}\begin{bmatrix}\text{-highT}\\\text{-lowT}\end{bmatrix}$, which is distinct from the other contours in Table 3.9. However, there are other problems with this alternative. Recall that non-syllabic segments are always [-stress]. Moreover, recall that the syllables which undergo sandhi must be stressed and must occur before a <u>stressed</u> syllable. If the syllabic vowel is deleted, then one must first have rules apply which specify the next vowel as [+syl] and [+stress] , if the syllable was stressed. This must be done irrespective of whether the other tone rules precede or follow this set of rules, for the stressed quality of the syllable must be retained. With either alternative, we see that if the extra-long forms are considered lexical, we must have a number of adjustment rules in the phonology just for these forms, none of which are motivated.

If the sandhi is considered to go in the opposite direction, then none of the problems concerning the canonical form of the syllable or the phonological rules arise. Therefore, let us now assume that the contours in Table 3.10 are lexical. The following rules will then express the alternations with the corresponding contours in Table 3.9.

(47). $$\emptyset \rightarrow \begin{bmatrix} \text{V} \\ \alpha F_i \\ \text{+lowT} \\ \text{-syl} \end{bmatrix} \Big/ \begin{bmatrix} \text{V} \\ \alpha F_i \\ \text{+highT} \\ \text{+stress} \end{bmatrix} \;\_\_\_\; \begin{bmatrix} \text{+son} \\ \text{-highT} \end{bmatrix}$$

where $\{F_i\}$ is the set of segmental features other than the pitch features.

(48). $$\begin{bmatrix} \text{+son} \\ \alpha\text{highT} \end{bmatrix} \rightarrow [-\alpha\,\text{highT}] \Big/ \left\{ \begin{array}{ll} \underset{\substack{\text{lowT} \\ \text{+stress}}}{\_\_\_\_} \begin{bmatrix} \alpha\,\text{lowT} \\ \text{+son} \end{bmatrix} ([\text{+son}]) & \text{(a)} \\ \begin{bmatrix} \text{-lowT} \\ \langle -\text{highT} \rangle \\ \text{+stress} \end{bmatrix} \overline{\langle \text{+lowT} \rangle} & \text{(b)} \end{array} \right.$$

Both (47) and (48) apply in the environment

$$/ \_\_ \left\{ \begin{array}{l}]_{\text{phon. phrase}} \\ C \begin{bmatrix} V \\ -\text{stress} \end{bmatrix} \end{array} \right.$$

. That is, the sandhi takes place whenever the syllable is stressed and occurs before a phrase boundary, or before an unstressed syllable, exactly as in rule (39) for Amoy. Rule (47) inserts a low toned, non-syllabic vowel in syllables with a high falling contour whenever they occur under the conditions stated above. Thus, (47), when applied to the sequence $\begin{bmatrix} V_1 \\ +\text{highT} \\ +\text{stress} \end{bmatrix} \begin{bmatrix} +\text{son} \\ -\text{highT} \end{bmatrix}$, will give the sequence $\begin{bmatrix} V_1 \\ +\text{highT} \\ +\text{stress} \end{bmatrix} \begin{bmatrix} V_1 \\ +\text{lowT} \\ -\text{syl} \end{bmatrix} \begin{bmatrix} +\text{son} \\ -\text{highT} \end{bmatrix}$, or in other words, the syllable is lengthened (˥˩ → ˥˩˨). Rule (48a) applies to all stressed syllables with a pitch contour where the first segment has the same value for [highT] and the opposite value for [lowT] as the second segment has for [lowT] . It therefore applies to the first segment in contours (2), (3), and (4) in Table 3.10, and assigns the opposite value of the feature [highT] . (48b) applies to the second sonorant in sequences which begin with [-highT] . The angle bracket notation states that if the preceding segment is also [-highT] , then the segment under consideration must also be [+lowT] . Notice that while (48a) applies to the contour resulting from the application of (47), (48b) does not. That is, the environment of (48a) permits the syllabic vowel to be followed by two sonorants before the phrase boundary or the first consonant of the unstressed syllable. The environment of (48b) will only permit one sonorant to follow the syllabic. Applying (48) to the contours in Table 3.10 and the contour resulting from application of (47), we get the following derivations.

(49). ˧ (47), (48) inapplicable.

(50). ˦˥ by 48a → ˥ by 48b → ˥˧ (47) inapplicable.

(51). ˥˧ by 47 → ˥˩˧ by 48a → ˧˩˥ (48b) inapplicable.

(52). ˩˥ by 48a → ˥˧ by 48b → ˥ (47) inapplicable.

(53). ˥˩ by 48b → ˧˥ (47), (48a) inapplicable.

(54). ˩ (47), (48) inapplicable.

Clearly, the derived contours are exactly those in Table 3.9.

We stated earlier that motivation for considering the underlying contours of Amoy to be those in Table 3.8, rather than those in Table 3.5, would arise from consideration of sandhi in Chaochow. Let us now proceed with that discussion. Amoy and Chaochow are considered to be very closely related dialects of Chinese. Amoy is thought to be the older of the two, because its consonant system is closer to that reconstructed for Ancient Chinese. That is, Amoy still has an initial /d/ and a final /n/ and /t/ in syllables, while Chaochow does not.[29] Nevertheless, the dialects are very closely related, much more so than, say, Amoy is to Fuchow, another Fukien dialect. The relation between the consonants and vowels of the one dialect to the other can be quite easily stated. Consequently, one would expect that one could also easily state the relations between the tone systems. It is generally assumed that the relations between languages (and hence, dialects) exist between underlying forms and not between derived forms. Thus, one would expect to see a close relationship between the underlying tone system of Amoy and that of Chaochow. We have argued above that the underlying contours for Chaochow must be those in Table 3.10. The correspondences between the Chaochow underlying contours as given in Table 3.10, and the Amoy contours when the underlying forms are considered to be those in Table 3.5, the isolation contours, are given in Table 3.11 below. The correspondences when the lexical contours for Amoy are considered to be those in Table 3.8, the non-phrase final contours, are given in Table 3.12 below.

TABLE 3.11

Chaochow ~ Amoy

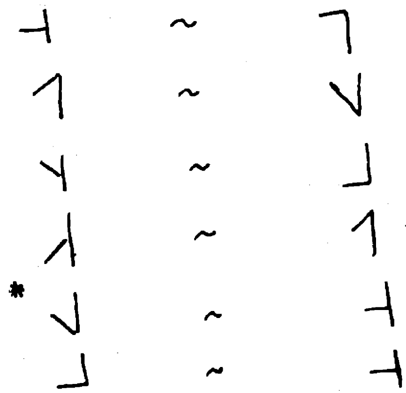
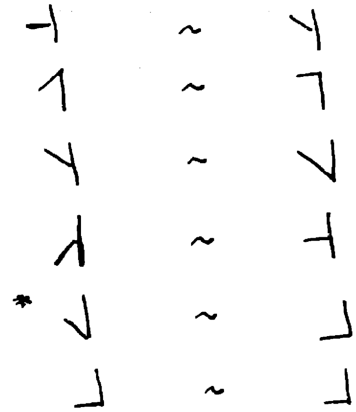


TABLE 3.12

Chaochow ~ Amoy



* Some of the Amoy syllables with a combination low tone (isolation mid tone) became low falling tone syllables in Chaochow rather than low level tone syllables. We do not know if there is any regularity characterizing this subset.

Looking at Tables 3.11 and 3.12, we see that the relations between Chaochow and Amoy tones are much clearer in Table 3.12, where the Amoy combination tones, or the non-phrase final contours, are considered lexical. We therefore feel that this gives additional motivation for considering the Amoy isolation tones to be derived and for considering the lexical contours to be those in

Table 3.8.

3.5 Amoy and Chaochow show that rules which affect the pitch features of one segment at a time are quite common. This kind of sandhi, one will recall, is the type we claim to be most frequently encountered. With rules of this type, one clearly has no need of bisegmental rules. Thus, we will now discuss the tone sandhi in Lungtu where there seems to be a case for bisegmental rules and the use of sequential marking conventions in the phonology.

Lungtu is a Fukien dialect spoken in the Chungshan district which is south of Canton and immediately north of Macao. Because the speakers of this dialect immigrated from Fukien to Chungshan hundreds of years ago, the language has been much influenced by the surrounding Cantonese dialects. Egerod, in his study, The Lungtu Dialect, describes the isolation tones as the following:[30] a high level tone (˥), a mid level tone (˧), a low level tone (˩), a mid to high rising tone (˧˥), and a mid to low falling tone (˧˩). These five contours only occur with syllables not closed by an obstruent. Syllables closed with an obstruent occur only with a short high or low tone. With open (non-obstruent closed) syllables, the following sandhi takes place: syllables with an isolation level high tone show a mid level tone in non-phrase final position; syllables with an isolation low level tone show a high level tone in non-phrase final position.

Unlike with Amoy and Chaochow, we have found no reason to believe that the isolation contours are not the lexical contours in Lungtu. Therefore, we shall assume that the sandhi rules derive the non-phrase final contours. We must now formulate a rule or rules to account for the alternations ˥ → ˧ and ˩ → ˥ , which will not affect ˧ , ˧˥ , or ˥˧ . The rules must apply whenever the syllables are in non-phrase final position. One could express these alternations using only rules which affect one segment. Such rules could be the following:

(55). $$\begin{bmatrix}+\text{son}\\ \alpha\,\text{lowT}\end{bmatrix} \rightarrow \begin{bmatrix}\alpha\,\text{highT}\end{bmatrix} / \begin{bmatrix}+\text{syl}\\ -\alpha\,\text{highT}\\ \alpha\,\text{lowT}\end{bmatrix} \_\_\_\_ \; (\text{CVV})_1 \Big]_{\text{phon.p.}}$$

(56). $$\begin{bmatrix}+\text{son}\\ \alpha\,\text{lowT}\end{bmatrix} \quad \begin{bmatrix}\alpha\,\text{highT}\end{bmatrix} / \_\_\_\_ \begin{bmatrix}+\text{son}\end{bmatrix} \; (\text{CVV})_1 \Big]_{\text{phon.ph.}}$$

Rule (55) will only apply to contours where the specification of the first segment for the feature [highT] is opposite to that for [lowT]. Thus, it does not apply to the lexical mid tone, the rising tone, or the falling tone. Rule (56) states that the specification for [lowT] becomes that for [highT] for all sonorants followed by another sonorant. Both rules, of course, only apply to syllables which precede one or more syllables within a phonological phrase. The rules apply to the contours above to give the following derivations:

(57). ˥ $\xrightarrow{\text{by 55}}$ ˥˩˥ $\xrightarrow{\text{by 56}}$ ˧

(58). ˩ $\xrightarrow{\text{by 55}}$ ˩˥˩ $\xrightarrow{\text{by 56}}$ ˥

The other three contours are not affected by (55) and (56) applies vacuously to them. Clearly, rules dealing with only one segment can be formulated. These rules are cumbersome, however, and do not reflect the fact that the level tones are alternating with level tones. If we make use of the bisegmental rules suggested earlier, we could express the sandhi as:

$$(59).\quad \overset{1}{\begin{bmatrix} +\text{syl} \\ \alpha\text{lowT} \\ -\alpha\text{highT} \end{bmatrix}} \; \overset{2}{[+\text{son}]} \qquad \overset{1}{[\alpha\text{highT}]} \quad \overset{2}{} \;\; / \underline{\quad}(\text{CVV})_1 \Big]_{\text{p.p.}}$$

(59) states that the contours which begin with a sonorant which has the opposite specification for [highT] as for [lowT] becomes a contour where the derived specification for [highT] is the one originally given for [lowT] . Moreover, the fact that this is a bisegmental rule and that no features have been specified for 2 on the right side of the arrow indicates that the second segment is considered unmarked and its matrix will be specified by the marking conventions to give level contours. Thus (59) will apply and give ˥ → ˧ , and

˩ → ˥ . (59) is not only simpler than the set (55) and (56), but it also captures the fact that in this language, the alternations are only between level tones. The ultimate decision as to which formulation is correct will, of course, depend largely on whether more examples can be found in which a formulation such as (59) is not only possible, but necessary.

3.6 From the preceding examples of the Chinese dialects of Amoy, Chaochow, and Lungtu, we see that for these languages, there are two primary factors in addition to stress which condition sandhi. The first is the position of the syllable in a phonological phrase, that is, whether the syllable is phrase final or not. Thus, in Amoy and Chaochow, the sandhi rules apply only to those syllables which are phrase final and stressed, with one exception. In Lungtu, the converse is apparently true. Sandhi which is conditioned by the fact that a syllable is phrase final is not unique to the Sino-Tibetan languages. In Mezquital Otomí, a language of Mexico, a tone change occurs whenever a bisyllabic word occurs before a phrase juncture, so that the second syllable assumes a high pitch.[31] In Otomí, therefore, bisyllabic words only show two contrastive pitch contours in isolation, rather than the four which exist phrase medially.

Traditionally, the isolation contours have always been accepted as the lexical contours for Chinese dialects. Thus, sandhi has always been said to apply to syllables which were no phrase final. However, as we have seen from the examples of Chaochow and Mandarin, to accept this view entails both complications in the lexicon and in the phonology because of the extra-long syllables which exist in both dialects. We have shown that if one considers the extra-long forms to be derived, then both the lexicon and the phonology are simplified. These simplifications, however, can only be achieved if one can state that in Chinese dialects, the environment for tone changes can be that a syllable is in phrase final position. We feel there is ample justification to consider this to be the case. One observation which justifies the statement is that then one can avoid listing the extra-long syllables in the lexicon, for they seem to only occur under unique conditions, specifically, when they are phrase final. This is true in Mandarin and Chaochow, with one minor exception, and apparently it is also true in Fuchow, which is reported to have both

˧˩˧ ($\begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix}\begin{bmatrix}\text{-highT}\\ \text{+lowT}\end{bmatrix}\begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix}$) and ˧˥˧

($\begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix}\begin{bmatrix}\text{+highT}\\ \text{-lowT}\end{bmatrix}\begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix}$). In Fuchow, as in Mandarin and Chaochow, these contours only occur phrase finally. Since the marking conventions for segmental features, as well as for pitch features, will reflect such configurations as costly, because a lexical entry with three contiguous sonorants, not including the prevocalic glide, is a highly marked entry, surely one would prefer to derive such forms whenever possible.

The second factor which conditions sandhi is the presence of a particular tone contour on the syllable immediately following the one under consideration. This kind of sandhi occurs in Mandarin. Although we have given only this one example of a language in which this kind of sandhi condition occurs, it is not an uncommon type of condition. It has been reported as the conditioning environment for sandhi in various of the Mexican Indian languages, for example, in Ayutla Mixtec[32], and in various of the African languages, for example, Ewe.[33]

On the other hand, sandhi conditioned by some feature of the initial consonant of the following syllable does not seem to occur. We will discuss in chapter IV an instance in which the tone of a vowel is affected by the following consonant, but this will be a case where the consonant closes the syllable. Although sandhi not being conditioned by initial consonants of following syllables seems to be a fact, the form of the rules expressing sandhi of the type found in Mandarin would seem to indicate sandhi conditioned by the pitch contour of the following syllable is more costly, and hence, less expected. That is, recall that the sandhi rule for Mandarin was:

$$(60).\quad \begin{bmatrix}+\text{son}\\+\text{lowT}\end{bmatrix} \rightarrow [+\text{highT}] / \begin{bmatrix}\text{V}\\+\text{lowT}\end{bmatrix} \_\_\_ (\#\#)C_o(G)\begin{bmatrix}\text{V}\\+\text{lowT}\end{bmatrix}$$

The initial consonant or glide of the next syllable had to be mentioned in the environment since the pitch features are specified as + or - for all segments, including consonants. However, it is not the inherent pitch of the consonant or glide which is the relevant environment, but the pitch specification of the following syllabic nucleus. In contrast, a rule of the type:

$$(61).\quad [+\text{lowT}] \rightarrow [+\text{highT}] / \underset{+\text{son}}{\_\_\_} [+\text{anterior}]$$

is evaluated as less expensive than (60). Yet, to our knowledge, tone changes are never conditioned by consonantal features such as [anterior]. To remedy this incorrect evaluation of these rules, we suggest that the evaluation of rules which affect the prosodic features, namely, the pitch features presented here, and the feature [stress], ignore the extraneous consonants and prevocalic glides which may be specified in the environment. That is, the features which specify these segments should not be included in the "feature count". To reflect the improbability of rules such as (61), we can only suggest that the evaluation of a rule which affects the prosodic features and which mentions as the environment a feature which is <u>only</u> a consonantal feature be given some great cost. By "<u>only</u> a consonantal feature" we mean those features for which the + and - specifications can both occur only if the segment is also [+cons] . Hence, for example, [nasal] and [voiced] are not in this set, but [anterior] and [coronal] are.

CHAPTER III: FOOTNOTES

1. We have resorted to calling these features [high tone] and [low tone] rather than [high] and [low] due to the unfortunate homonymity of these latter names with the segmental features, [high] and [low].

2. Personal communication, Eunice Pike. I must thank Miss Pike for giving me a copy of a tape with data from Kent Wistrand.

3. Cheng, Mandarin Phonology, especially chapter IV.

4. Hockett, "Peiping phonology".

5. Martin, "Problems of hierarchy and indeterminacy in Mandarin phonology".

6. C.C. Cheng, in his thesis, Mandarin Phonology, argues at length that there are only four underlying pitch contours in Mandarin, and that lexical neutral tone syllables are specified with one of these four citation contours and with the diacritic, [+neutral]. Phonological rules, Cheng states, would interpret this diacritic as [-stress] and all syllables with the specification [+neutral] would be subject to the neutral tone realization rules.

I do not agree with Cheng. However, I would now like to briefly state his arguments and present my objections to it. Cheng's principal reason for arguing that the lexical neutral tone syllables are specified with one of the four citation contours arises from his feeling that these syllables must be closely related in the synchronic phonology to monosyllabic lexical items which are specified with one of the four citation contours. Thus, he argues, for example, that the suffix /-tz/ 子 , a nominal suffix which adds no meaning to a noun, is specified lexically with the same pitch contour as the item /tzyy/ 子 <u>child</u>. (I am once again using the National Romanization system to transcribe the characters.) It is true that the same graph is used in the writing system to transcribe both items. However, that cannot be an argument that the lexical representation of the two items must be identical. There are many instances of lexical items which are neither semantically, syntactically, nor phonetically similar, and yet are transcribed with the same graph. One example of this is the graph 行 . Two of the readings given to this character are romanized as /shyng/ and /harng/ (phonetically, /ʂiŋ / and /haŋ /). With the first pronunciation alone, the item has several meanings, among which are: *to walk*, *to go*, and *to perform* (the latter meaning only occurs if the syllable occurs in a complex verb phrase). With the second pronunciation, the item again has several different meanings, among which are: *a row* (auxilary noun), and *business house*. The various forms transcribed with the graph 行 probably were once etymologically related. However, there is no synchronic reason to think that they are. Consequently, there is no reason to believe that the lexical specification of any one of these forms bears any resemblance to the others. Likewise, the suffix /-tz/ and the noun /tzyy/ were probably related historically, but there is no reason to think that they are related synchronically outside of the fact that they are transcribed with the same graph.

There are two principal objections to specifying the lexical neutral tone syllables with one of the four citation tones. The first is simply that there are a number of lexical neutral tone syllables which are part of bisyllabic lexical items and for which it is difficult to find a monosyllabic lexical item which can be said to be related, such that the pitch specification can be determined for the neutral tone syllable. For example, in the bisyllabic lexical items, /dong shi/東西 <u>thing</u>, and /shyng lii/行李 <u>luggage</u>, the second syllable in each item is a lexical neutral tone syllable. However, if it is to be specified with the diacritic [+neutral] and must be specified with one of the four citation contours, then it is difficult to determine what the underlying contour for each of these syllables should be. That is, the noun /dong shi/ is transcribed with the graph 東, which by itself means <u>east</u>, and the graph 西, which by itself means <u>west</u>. Clearly the compound /dong shi/ does not derive its meaning from the separate meanings of the graphs. Thus, what monosyllabic lexical item can one associate with the second syllable in /dong shi/ to determine what its lexical pitch specification should be? Likewise, the meaning of /shyng lii/ is clearly not a derivative of the meanings of the graphs used to transcribe it. Cheng admits that he cannot determine what the lexical contour of the second syllable in these items should be. Such items, however, are not rare. One could simply state that the tone contours of the second syllable is that of the reading given to the graph used to transcribe it, but there is no justification for this. Moreover, there are some lexical bisyllabic items which occur in the spoken language which have the stress pattern, "stressed - unstressed", which do not occur in the written language. One such item that Chao cites is /tzao chin/ morning (Chao, Mandarin Primer, pg. 168, note 13). Here not only can one not determine what the lexical contour must be for the second syllable, because there is no phonetically similar monosyllabic lexical item which is also semantically related, but one cannot even state that the contour would be that of the reading pronunciation of the graph used to transcribe the second syllable, for the item is not written, and hence, there is no graph.

The second objection to specifying the neutral tone syllables with a citation contour arises from the fact that then the phonology must be complicated. There are bisyllabic nouns which appear to be reduplicated nouns, but which must be listed in the lexicon for various reasons, rather than derived. Such items include /jiee jiee/姐姐 elder sister, and /ge ge/哥哥 elder brother. For the most part, these are relational terms and the list is quite small. Cheng feels that because the graphs for both syllables in each item are identical, then the feature matrices of both syllables should also be identical. For the item /jiee jiee/, this would mean that the pitch contour of both syllables is [+lowT] [+lowT] in our system. If the pitch specification of both syllables is that given above, then the sandhi rule which changes a low level tone to a rising tone whenever it occurs before another low level tone should apply to the contour of the first syllable in /jiee jiee/. However, the phonetic contour of the first syllable in this item is not a rising contour, but a low level contour. Cheng blocks the sandhi rule from applying to this and similar items, by stating a special rule which erases all internal boundaries in these items. The rule is somewhat <u>ad hoc</u> for it states that these items are in some arbitrary way significantly different from /dong shi/, for example, which, for Cheng, apparently does contain internal boundaries. If the second syllable of /jiee jiee/ is not specified as [+lowT] [+lowT], on the other hand, but rather as $\begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix} \begin{bmatrix}\text{-highT}\\ \text{-lowT}\end{bmatrix}$, then no special adjustment rules need apply to this form. The fact that apart from the specification of the pitch features, the matrices of the two syllables are identical, should be sufficient to relate the two syllables if indeed they are related.

In summary, my objections to specifying the lexical neutral tone syllables with one of the four citation contours in Cheng's manner arises from the fact that it is often difficult, if not impossible, to determine which contour should be specified, and the fact that this solution entails unnecessary complications in the phonology. I feel that the alternative, specifying a fifth contour, is therefore preferable, in particular, since the fifth contour does not add to the complexity of the lexicon since it will be seen that it is the unmarked contour.

7. See Li, "Tone Perception".

8. The slashes (//), as usual, generally enclose the romanized form of the syllables. The times when a phonetic representation is given between slashes are, I think, obvious.

9. Kratochvíl, pg. 388, for example.

10. Chao, <u>Mandarin Primer</u>, pg. 90.

11. Wang and Li, "Tone 3 in Pekinese".

12. Dreher and Lee, "Instrumental Investigation of Single and Paired Mandarin Tonemes".

13. That is, the sandhi rule not only applies in phrases such as /woo mae bii/, where each syllable is a phonological word, but it must also apply to morphologically and syntactically derived bisyllabic words such as /tzoou tzoou/走 走 to take a walk. In /tzoou tzoou/, both syllables are specified as [+lowT][+lowT] . However, the second syllable is unstressed, and therefore a clitic. Nevertheless, the sandhi rule must still apply to the first syllable to give a rising contour. Although I have only stated that the boundary erasure rule must precede (12), since I believe it is not a Mandarin-particular rule, but a universal convention which applies to all clitics, I have assumed that the boundaries will be erased before rule (5) applies. Hence, the particular formulation of the environment of (5).

14. Egerod, The Lungtu Dialect, pp. 27-28.

15. Of course, there will be some diphthongs which are less "natural" than others. For example, one

might wish to say that /ai/ is more "natural" than /ae/. This would simply imply that the sequential marking conventions for vowel features are not as simple as the ones we have proposed for the pitch features.

16. Personal communication, LaRaw Maran.

17. _SPE_, pp. 419-35.

18. Bodman, _Spoken Amoy Hokkien_, vol. 1, pp. 38-41, and pp. 84-86.

19. Lo Ch'ang P'ei, _Phonetics and Phonology of the Amoy_ pg. 22.

20. _Hanyu Fangyan Cihui_, pg. 14.

21. Bodman, _Spoken Amoy Hokkien_, pp. 39-40.

22. Bodman, pg. 39.

23. Wang, "Phonological Features of Tone", pg. 105.

24. Wang, pg. 103.

25. Wang, table 1, pg 97.

26. See, for example, Lo Ch'ang P'ei, pg. 22, and _Hanyu Fangyan Cihui_, pg. 14.

27. Li Yong-ming, Phonology of Chaochow, pg. 14.

28. Hanyu Fangyan Cihui, pg. 15.

29. Li Yong-ming discusses the relationship of the two dialects on pp. 1 ff.

30. Egerod, The Lungtu Dialect, pg. 49.

31. personal communication, Ethel Wallis.

32. Pike, Eunice and Pankraz, Leo, "Phonology and Morphotonemics of Ayutla Mixtec".

33. Smith, Neil, "Tone in Ewe".

CHAPTER IV: THE ARTICULATORY CORRELATES

4.0 In this chapter, we will discuss three topics. The first will be the articulatory correlates to the features for tone presented in chapter III. The second will concern other linguistic factors which can affect the pitch contour of a syllable or word. In this section, we will attempt to account for the source of some of the phonetic dynamic tones which have been reported as occurring on syllables with short vowels. Among the topics to be covered, there will be the effect of the consonants on the pitch contour of the vowel in syllables with a CVC structure. The last section will deal with hypotheses as to how a language can become tonal in the sense of having pitch distinctions in the lexicon. Obviously in this last section, we cannot give a complete or final answer to the question of how tones arise, but we will see from material which will be given on Jingpho Kachin, how Haudricourt's hypothesis[1] on the origin of tones in Chinese and Vietnamese appears to gain support from the material presented in all three sections of this chapter.

4.1.0 In chapter III, we presented the set of pitch features, [high tone] , [low tone] , and [modify] . These three features, along with the feature [stress] , are the distinctive features of prosody. In SPE, Chomsky and Halle propose a third category of prosodic features, namely, length. So far as our studies are concerned, we have been unable to find evidence to support the hypothesis that the feature [long] is needed in phonology, such that vowels with the feature [+long] must be considered as different from geminate vowels. In all of the languages examined here, where the pitch features represent the primary word level prosodic phenomenon, length of vowels has always been associated with a geminate vowel, where the existence of the latter has been motivated by considerations such as symmetry of syllable structure, and conformity of long vowels with diphthongs and vowel plus nasal clusters with respect to such rules as the sonorant deletion rule for Mandarin unstressed syllables. We therefore have not found any justification for positing the existence of a third phonological prosodic feature category of length.

We must make it clear that the abstract representation of long vowels as geminate clusters does not imply that there is an absolute notion of length in a given language any more than the traditional feature [long] did. That is, there can always be phonetic variation in the interpretation of the length of a geminate vowel. All that we claim is that a geminate vowel in a given language, and in a given environment, will be articulated with longer duration than a single vowel in the same environment. Thus, an unstressed geminate cluster in some language may be, for example, 2.0 times longer than an unstressed simple vowel. In the same language, the primary stressed vowels will have the same ratio of duration, but the stressed geminate cluster may also be, say, 1.5 times as long as the unstressed geminate cluster. These length variations, we feel, do not necessitate the recognition of a third category of prosodic features, but rather, they imply that the phonetic correlates to [stress] be recognized as more complex than simple intensity.

4.1.1 There has been some question in the past as to the relation of "stress" and "high pitch". Because in English, as well as in some other languages, the syllable with primary stress is generally manifested with both higher intensity and a higher fundamental frequency (pitch), some question arose as to the independence of the feature [stress] from the pitch features. In addition to the fact that there was a correlation between intensity and pitch in languages like English, there also appeared to be no such phenomenon as "stress", in terms of intensity, reported in languages with lexical tone. Because of these two considerations, some linguists have wondered whether the feature [stress] and the pitch features had the same property as palatalization and pharyngealization, for example, namely, that these latter two articulations are mutually exclusive on any given segment. In fact, some linguists have even wondered if stress and pitch were not separate manifestations of the same articulatory mechanisms.

In this work, we have shown that [stress] is indeed distinct from the pitch features. This follows directly from chapter I, and from the material in chapter II on Mandarin, where the phonetically distinctive characteristic of some bisyllabic pairs was the stress, or intensity, pattern. [-stress] is, in fact, the characteristic property of "neutral" tone syllables. This has therefore led us to conclude that stress and pitch are independent features and are produced by different mechanisms. The primary reason for the confusion about the status of these two features has been the result of the fact that in languages with lexical tone, either the language is primarily monosyllabic in its vocabulary, like the Oriental languages, and hence, stress patterns in polysyllables are not studied since the polysyllabic items themselves are often not mentioned, or there is a regular association of primary stress with the high tone syllables, as in Bambara. In the latter case, the languages generally only have two contrastive pitches, high and non-high. Above all, however, one must remember that stress is simply not discussed for most languages with lexical tone. There are, of course, languages in which the value of either [high tone] (or [low tone]) or [stress] is predictable from the other. Such is the case in English, where high tone generally occurs with primary stress. In other cases, however, such as in Mandarin, tone is lexical, and primary stress is not correlated with any particular tone configuration. We shall discuss later how stress can interact with pitch even when the two sets of features function independently.

We cannot give a precise definition of stress, but can only state that among the most common phonetic correlates to [+stress] , one finds higher intensity, raised pitch, and increased duration of sonorants. A more detailed discussion of the acoustic correlates of [stress] can be found in Lieberman's study, Intonation, Perception, and Language.

4.1.2 The features [high tone] and [low tone] characterize the position of the glottis, the length and thickness of the vocal cords, and the tension of the vocal cords. As we shall see, these two features can be used to characterize a raised and lowered position of the glottis relative to some neutral rest position in a completely analogous manner to the segmental features [high] and [low], which refer to the position of the body of the tongue relative to some rest position. The neutral position for the body of the tongue is described in SPE as "raised and fronted, approximately the configuration found in the vowel /e/ in English bed."[2] In a similar manner, we assume the neutral position of the glottis, the thickness and tension of the vocal cords, and the length of the vocal cords, to be that found in the production of the "natural pitch" for any given speaker. Zemlin describes this "natural pitch" as somewhere around "middle C" (C_4) for female speakers, and somewhere about an octave lower, C_3, for male speakers.[3] From various studies on pitch and pitch variability, this "natural pitch" is found to be approximately one quarter up the total singing range, including falsetto.[4] When any speaker is engaged in relaxed phonation, his voice may produce a variety of tones ranging in pitch over approximately two octaves. Generally, however, these pitches are distributed so that a central tendency is observed. This mode, or central tendency, is what we call the "natural pitch" or "habitual level". Obviously, a speaker can produce pitches which are higher or lower than this habitual level. What we must now specify is the nature of the mechanism actually employed in normal phonation as opposed to the mechanisms which could theoretically be employed.

If the vocal cords behaved like a stringed instrument, then a simple <u>decrease</u> in vocal cord length should produce a higher pitch. However, they do not, and observations of laryngeal behavior show that the converse appears to be true. That is, the length of the vocal cords is <u>increased</u> as the pitch increases (gets higher). The pitch raising mechanism must therefore involve a change in the mass and/or tension of the vocal cords. Indeed, as the vocal cords lengthen, laboratory observations have noted a simultaneous decrease in the cross-sectional area (mass). However, the observed decrease in mass in the production of high pitch has never been found to be as great as would be necessary if decrease in mass were solely the factor involved in pitch raising. Consequently, one must conclude that vocal cord tension plays an important part in pitch change. It is an open question, and for us, an irrelevant one, whether or not the tension change is the only significant operation with the length and mass changes resulting automatically. Suffice it to say that the increase in pitch from the habitual level is correlated with an increase in vocal cord tension, an increase in vocal cord length, and a decrease in vocal cord thickness. In addition to these three correlations, recent laboratory studies have shown that an upward movement of the glottis occurs with the production of high pitch, and a downward movement with low pitch.[5] This is by no means an observation limited to one study, for Zemlin writes, "It is a rather common occurrence for the larynx to rise during the production of high pitched tones and lower during production of low tones."[6] This movement is accomplished through use of the extrinsic laryngeal musculature, and is particularly noted in the production of tones at the extreme ends of the pitch range. When pitch is used as a feature in a language independent of other segmental features such as stress, the differences between the tones used tend to be gross rather than slight. Therefore, it is not surprising that the extrinsic musculature also plays a role, for the tones used in languages must be at the extreme ends of the pitch range. This becomes even more evident when one considers that pitch differences must be maintained in speech, where both intonation and stress may affect any particular syllable greatly, and hence destroy pitch differences if they are only slight. In general, the extrinsic muscles will raise the larynx above the rest position during high tone production. This is not always the case, however, for other factors may render this impossible. For example, in Igbo, when a high toned vowel follows an imploded consonant, the glottis is not observed to rise beyond the neutral position. However, this is only reasonable, for in the production of the implosive, the glottis is greatly lowered; with our informant, the Adam's apple disappeared below his collar.[7] When a high toned vowel followed, one saw that the glottis rose, but because of the extremely low starting position due to the transition from the implosive, the maximum height attained was approximately the neutral position. Nevertheless, the glottis did rise. The various displacements of the vocal cords with respect to mass, length, and tension, and the displacement of the glottis relative to the neutral position, are characterized using the features [high tone] and [low tone] in the following ways.

High tone - non-high tone High tone sounds are those produced with the vocal cords longer, tenser, and less thick than in the neutral position. In addition, the glottis is raised above the neutral position. Non-high tone sounds are produced without the raising of the glottis, and without lengthening or tensing or decreasing the mass of the vocal cords.[8]

Low tone - non-low tone Low tone sounds are produced with the vocal cords shorter, laxer, and thicker than in the neutral position. In addition, the glottis is lowered below the neutral position. Non-low tone sounds are produced without the lowering of the glottis, and without the shortening, laxing, and thickening of the vocal cords.

The movement of the glottis above and below the neutral position has been observed and measured externally for a Jingpho speaker, LaRaw Maran. The deviation is in the vicinity of 1 to 2 cm. above or below the rest position when observed in the production of high and low tones. Jingpho Kachin has three pitch heights and no movement of the glottis was detected in the production of the mid tone.

We see from these characterizations why no segment can be $\begin{bmatrix}\text{+high tone}\\ \text{+low tone}\end{bmatrix}$, for it is impossible to both raise and lower the glottis, to both lengthen and shorten, thicken and thin, or tense and lax the vocal cords. Clearly, all the sonorant segments can theoretically be produced with a high pitch, a low pitch, or a non-high, non-low pitch (the "natural" pitch). Indeed, the only constraints placed on sonorants in the marking conventions in chapter III were in the statement specifying word initial non-syllabic sonorants as non-contrastive in pitch. In rapid speech, it is possible that a word initial nasal or glide will assume the pitch of the immediately preceding sonorant, if the glide or nasal is phrase medial, and follows a word final sonorant. However, in careful speech, where word junctures are pronounced, and in isolated utterances, there is no evidence showing that word initial sonorants carry contrastive pitch.

Non-sonorant consonant sounds, unlike sonorants, cannot be produced freely with a high, mid, or low pitch. In the last chapter, we noted that traditionally, voiced consonants were associated with low tone, and voiceless consonants with high tone. There it was suggested that the inherent pitch of the consonants was a consequence of the vocal cord configuration. Voiceless obstruents are produced with a spread vocal cord configuration. Because of the spread, it is not improbable that the vocal cord tension is then greater than that in the production of voiced obstruents. If an increase in vocal cord tension automatically results in vocal cord lengthening and thinning, then three of the requirements for production of high tone are fulfilled by voiceless obstruents. LaRaw Maran has also observed a slight raising of the glottis in the production of voiceless consonants and a lowering with voiced consonants.[9] It remains to be seen whether this is always the case, and in particular, whether this motion, if it always exists, is an automatic consequence of other factors in the production of voiced and voiceless obstruents. One may well ask why there are no obstruents with a non-high, non-low pitch. If nasals are considered as obstruents, then the nasals fill this gap. However, since this particular characterization of nasals is totally out of the question, we cannot cite them as an example. Instead, we must state that the answer is now unknown.

Having defined the features [high tone] and [low tone] , we may ask if these features can be used to characterize any sets of consonants. Recall that high tone sounds are produced with a raised glottis, while low tone sounds are produced with a lowered glottis. A raised glottis, however, has also been found to be the characteristic property of ejectives, while implosives are produced with a lowered glottis. Thus, Ladefoged states, "The glottalic airstream mechanism is not used in the well known European languages (except as a stylistic variant); but in West African languages it is involved in the production of two kinds of sounds: ejectives, in which the closed glottis moves rapidly upwards and compresses the air behind an articulatory closure; and implosives, in which there is a downward movement of the closed glottis, which tends to cause a lowering of the pressure behind the oral closure."[10] We see, therefore, that the mechanism involved in producing high tone sounds and low tone sounds is also found in the production of certain consonants. However, since we have previously stated that movement of the glottis has been observed in the production of ordinary obstruents, clearly the position of the glottis is not enough to differentiate, for example, between normal voiced plosives and implosives. Ladefoged states, "A voiced plosive may have no downward movement of the glottis at all; but quite often (even in an ordinary English b) there is a lowering of the larynx ......Few people would call such a sound an implosive; but the mechanism is essentially the same as in Kalabari ɓ, differing only in the rate and degree of larynx movement."[11] He then suggests that the voiced plosive /b/ in Igbo might be distinguished from the implosive /ɓ/ by the fact that the latter is "velarized as well as usually involving lowering of the glottis."[12] We see, therefore, that the mechanism involved in producing high and low tone sonorants is also found in the production of ejectives and implosives, although apparently the features [high tone] and [low tone] cannot by themselves completely differentiate ejectives and implosives from normal released stops.

4.1.3 The features [high tone] and [low tone] account for three pitch heights, but what about the languages reported to make a distinction of five pitch levels? Because we have had no opportunity to study any of these languages first hand, we can only suggest the nature of the mechanism involved in the production of the two intermediate levels. Given that the mechanism for producing high pitch and low pitch is as discussed in the preceding section, it is difficult to imagine that the intermediate pitches are produced using only this mechanism. To do so would involve incredible control of the laryngeal muscles. In fact, if one assumes that the differences between all five pitches is solely one of frequency, then one must assume that listeners can consistently detect five pure tones. This has not been supported by laboratory findings. Pollack reports in his article, "The Information of Elementary Auditory Displays", that,

> "When the frequency of a single tone is varied in equal-logarithmic steps in the range between 100 cps and 8000 cps (and when the level of the tone is randomly adjusted to reduce loudness cues), the amount of information transferred is about 2.3 bits per stimulus presentation. This is equivalent to perfect identification among only 5 tones. The information transferred, under the conditions of measurement employed, is reasonably invariant under wide variations in stimulus conditions."[13]

This does not mean, however, that when presented with exactly five tones, the listeners could correctly identify all of them. Rather, the listeners could identify five only when presented with <u>eight</u>. When presented with exactly five tones, ranging in frequency from 100 cps to 8000 cps, the information transferred ranged between 1.8 and 2.1 bits, or identification of three to four tones, never all five.[14] Pollack's results are interesting also in that they show that the range is largely irrelevant, that is, he tested identification of tones in ranges with the lower limit at 100 cps and the upper limit varying from 500 cps to 8000 cps, and found the results to be essentially the same. The information gained increased slightly as the range increased, but never exceeded the maximum of 2.3 bits. In the range which is the maximum for a natural speaking voice, 100 cps to 1000 cps, his results showed the information transferred to be only 1.9 bits, or identification of four tones, even though <u>eight</u> tones were presented. Other attempts to improve identification were made involving changes in the distribution of the tones and the number of tones employed in a series, but the results did not improve beyond the maximum 2.3 bits.

Because of the difficulties found with identification of five pitches differing only in frequency, we feel that in languages where there are five pitch levels reported, the tones are produced using more than one dimension of sound stimulus. What the second dimension is, we do not know, for that can only be determined from direct laboratory studies of such languages.

One highly plausible possibility is that the intermediate levels are produced in the same way as the high tone and low tone are, except that there is a change in the quality of the pitch through a change in the waveform. This modification can arise in many ways, such as through a change in the manner in which the vocal folds vibrate, or a change in the configuration of the pharynx, which would result in a change in the shape of the resonance chamber. The first method, change in the manner in which the vocal folds vibrate, is one of the several ways in which voice quality can differ from one speaker to another, along with size of vocal tract, length of vocal cords, etc. Both ways of modifying a pitch would yield a different waveform, or overtone pattern, but not necessarily entail a change in the fundamental frequency. Differences in overtone pattern could give rise to perceptual differences which would be equated to a simple difference in fundamental frequency. In all probability, the intermediate levels manifest both a waveform difference as well as a slight frequency difference, in relation to any contiguous non-modified high toned or non-modified low toned sonorant.

In chapter III we stated that no segment could be specified as $\begin{bmatrix} \text{-high tone} \\ \text{-low tone} \\ \text{+modify} \end{bmatrix}$. Why? Clearly it is physiologically possible to impose the modifications of the waveform on the vocal cord and glottis configuration of the mid tone. We base our decision to exclude this configuration as phonologically distinctive on the belief that variations of this type on the "natural pitch" would be too minute to be perceivable. Since the mid tone in particular is subject to great variations just in terms of frequency from one utterance to another for any given speaker, we do not feel that a waveform difference could be easily distinguished in the middle range, if at all. On the other hand, judging from the ease with which a listener can identify a singer on high or low notes, as opposed to middle-range notes, we feel that the deviation at the extreme points of the pitch range can be more readily perceived. In particular, if, as is probably the case, the intermediate levels show both a waveform and a frequency difference from the pure high and low tone sounds, and the frequency of the intermediate levels is not in the range permitted for a mid tone in a given environment, then it is obvious that the sounds produced do not satisfy the criterion for non-high tone, non-low tone sounds. That is, the mid-high level tone is generally described as higher pitched than the mid level tone. Consequently, we should find an increase in vocal cord tension and length, and a decrease in vocal cord thickness, or a [+high tone] sound. We therefore define the feature [modify] as follows:

<u>Modify - non-modify</u> A modified sound is one produced with the vocal cord and larynx configuration of either a high toned sound or a low toned sound and is one in which either the configuration of the pharynx has been changed, or the manner in which the vocal folds vibrate has been changed, or some other mechanism has been employed to change the "normal" waveform of the pitch produced. The normal waveform referred to is that associated with the pure high toned or pure low toned

sounds. A non-modified sound is produced without the change in waveform.

In the preceding, we stated that the intermediate levels probably differed both in the waveform and in fundamental frequency from the high and low pitches. Our reason for stating that there is probably a concomitant frequency change is due to the fact that although one can theoretically assume variations along a single parameter in an abstract model, in reality, very little can change in one part of the larynx which does not affect other parts. In particular, it is not improbable that a change in the configuration of the pharynx or in the manner in which the vocal folds vibrate should simultaneously affect the tension of the vocal cords.

For sonorant sounds, it appears that one can manipulate the larynx rather freely without interfering with the production of the sound. For non-sonorants, however, it is somewhat doubtful if the feature modify could ever apply. Clearly, for voiceless sounds, where the vocal cords are not vibrating, one cannot change the manner in which the vocal folds vibrate. With voiced obstruents, it is rather doubtful if the the manner in which the vocal cords vibrate can be adjusted, because of the other constraints imposed on the articulation of voiced obstruents. If the waveform adjustment results from a change in the configuration of the pharyngeal musculature, in all probability, the result is simply a pharyngealized consonant. All of this must, of course, remain highly speculative until more intensive laboratory studies can be conducted.

4.2.0 The preceding discussion has only been concerned with the articulatory mechanisms involved in producing a "pure" tone, or level pitch on a single sonorant. Contour tones on sonorant sequences will result when the phonetic instructions for pitch differ from one segment to the next. Now we must discuss the contour tones which are sometimes perceived on phonetically short vowels. Pitch contours on short vowels may be perceived as the result of stress and intonation mechanisms interacting with the pitch production mechanism, or as a result of the transition from a particular type of consonant to the vowel, or from the vowel to a particular type of consonant. In this latter case, we shall only speak of the effects of surrounding voiced and voiceless obstruents, although it is not obvious that this is the only feature of consonants which would have an effect on pitch.

If one could produce a sonorant sound without changing the subglottal pressure at any time during its articulation, then one would expect to always get a steady fundamental frequency curve. However, since phonation involves the passage of air in the production of sonorants, the air pressure will not always be a constant. In particular, under conditions of primary stress, one often observes a falling pitch phenomenon, especially if the primary stressed syllable is followed by an unstressed syllable. It is a well known fact that with a change in subglottal pressure, one can observe a proportionate change in pitch. If, as in English, there are no conflicting phonetic instructions for pitch in a sonorant, then this will mean that a primary stressed syllable is high toned. However, when an instruction for pitch already exists, the stress may only serve to raise the pitch of the syllable slightly.

4.2.1 We will not discuss all of the cases of short vowels affected by primary stress, but rather just one instance as an example. Quite often one will hear of low-toned short open syllables in a language being perceived with a falling contour when stressed and as level when unstressed. Such a phenomenon is reported to occur in Jemez, for example.[15] Clearly, since the contour is so obviously associated with stress, one wonders if in fact the stress has produced the contour. In particular, the contour is noticeable when the stressed syllable is not word final. The contour, then, is probably the result of the low pitch being raised during the peak of the intensity, the correlate of [stress] in Jemez, and then decreasing as the intensity decreases. Since the syllable is non-final and since succeeding syllables are not stressed, the intensity must decrease rapidly. Why does one not perceive the contour as rising-falling rather than just falling? The most plausible answer is that the intensity peak is not attained in the middle of the total vowel articulation, but rather nearer the beginning. Another is that one does not notice the rising intensity and pitch, that is, that the information transferred only comes from the portion of the vowel beginning with the intensity peak and that following it. This would not be unusual, since people do not apparently notice the "dips" in the frequency curves in Mandarin syllables which we have mentioned before. It is quite likely that acoustic analysis will show both a rise and a fall in the pitch contour for Jemez syllables, however, it is also likely that the degree of the rise is less than that of the fall, and that the duration of the rise is less than that of the fall.[16]

There are also instances when a stressed low tone is heard as lower than its unstressed allotone. This is the situation reported in Trique,[17] for example. However, in Trique, the high tones were also perceived as higher under stressed conditions. This "exaggeration" of the pitch range under stressed conditions has also been noted by Chao, in a discussion of stress in Mandarin.[18] It is possible that in these cases, an increase in intensity has served to emphasize the pitch levels and therefore a stressed vowel would appear "more" high (or low) pitched in these environments. This is similar to the situation with stressed vowels as opposed to unstressed vowels, where in many languages, the former are perceived as having "clearer", or "sharper" qualities than their unstressed correlates.

The whole problem of the interaction of pitch and stress is rather confused by the fact that it is still undertermined as to exactly what people perceive as stress. A rapid decrease in intensity, however, is associated with falling pitch in the unstressed syllables in Mandarin when these syllables are phrase final. Dreher and Lee report a 41 contour for unstressed syllables after a stressed high toned syllable (5 is high on this scale and 1 is low).[19] However, they only tested unstressed syllables in bisyllabic utterances where the unstressed syllable was phrase final. Under these conditions, one finds that the intensity of the unstressed syllable decreases rapidly, thus causing a sharp decrease in fundamental frequency. When the unstressed syllable is not phrase final, however, one does not find this pitch contour. One must therefore conclude that this change in intensity causes the contour on the Mandarin unstressed syllables.

The interaction of stress with pitch is in many ways analogous to the interaction of stress with the other sonorant features. Under conditions of primary stress, one often finds that there are more vowel distinctions than under [-stress] conditions. The same appears to be true for tones. In Trique, for example, primary stress occurs on the last syllable of a word (excluding clitics), and it is only on the last syllable that one finds all five pitch levels. Primary stress may also be the conditioning environment for tone sandhi as it is for the vowel shift in English. For example, the extra-long form in Mandarin only occurs if the syllable is both phrase final and stressed. Thus, if the final syllable has an underlying low tone, but has been assigned [-stress] by the stress placement rules, then the elongation does not occur.

Hence, in the prepositional phrase /u lii/屋裡 inside the room, where the preposition /lii/ has the lexical pitch specification [+lowT] [+lowT] , the syllable has been assigned [-stress] by the part of the stress placement rules concerning prepositional phrases. The syllable is then subject to the neutral tone realization rule rather than to the elongation rule. Likewise, in Amoy, the syllables affected by the sandhi rules discussed in chapter III must be [+stress] . It is therefore clear that stress can interact with tone changes in much the same way that it does with vowel changes. On the one hand, vowel reduction generally occurs when the vowel is [-stress] . Similarly, tones "reduce" when the syllabic sonorant is unstressed, as in the case of the Mandarin neutral tone syllables. On the other hand, the stressed tense vowels in English undergo vowel shift. Similarly, the stressed sonorant clusters in Amoy which are phrase final, or before an unstressed syllable, undergo "tone shift".

4.2.2 The effects of sentence intonation on tones has not been studied for the most part. In general, all that one can find about the effects of intonation on the tone of a word concerns the final syllable of a sentence. There it is a well noted phenomenon that the intonation of the sentence may modify the tone of the last syllable beyond recognition. Other than that, however, most writers claim that there is no such thing as "intonation" in a tone language except in as much as some syllables are more highly stressed than others inside of a phrase. S. Öhman, in a study of Swedish sentence and word intonation, has attempted to construct a model describing the interaction between the two, and points to some interesting possibilities for further study.[20] Unfortunately, in Swedish the prominence peaks for a phrase are assumed to occur over accented syllables, and hence, little modification of the pitch contour apparently occurs. What would be of interest to determine is the actual contour of, for example, a low toned syllable with 3 or 4 stress preceded by a falling tone syllable with secondary stress, and followed by a high toned syllable with primary stress on which also occurs the prominence peak, in Lieberman's terminology, of the breath group. Under these conditions, that is, of having 3 or 4 stress and surrounded by primary or secondary stressed syllables, it is reasonable to expect that even the short syllables will show a pitch contour. Thus, this is another possible source of short vowels with contour tone.

One rather interesting example of the interaction of sentence intonation with the tone of a syllable is reported by Neil Smith to occur in Ewe.[21] Smith states that whenever a sentence contains an overt marker of the interrogative, the sentence will terminate on a low pitch. Thus, if the last syllable of the sentence is high toned, it will be produced with a short falling tone. In other types of sentences, no such fall is observed. Apparently the intonation contour for sentences containing overt question markers is one which drops sharply at the end, thus causing falling tones to be perceived even in short syllables if they occur at the end of these sentences. The situation is not unlike one found in English with direct questions, where the last syllable of the sentence is always high pitched because the intonation contour for direct questions rises sharply at the end of the sentence. In English, it is apparently irrelevant whether the last word in a direct question has primary stress or not, for the pitch always rises on it. Thus Lieberman found that in the pair of sentences,

(1). Did Joe eat his soup?

(2). Did <u>Joe</u> eat his soup?

where <u>Joe</u> has the greatest stress in (2), <u>soup</u> always has the highest pitch and shows a rising contour.[22]

4.2.3 The last way in which short vowels can manifest contour tones which we will discuss is through the effect of the surrounding consonants. It is a well known fact that the voicing of consonants affects the pitch of surrounding vowels. If the consonant is not in the same syllable as the vowel, then the effect may not exist, or may be negligible. However, when the consonant either opens or closes the syllable in which the vowel occurs, then the effects become more marked. Voiced consonants can depress the pitch, while voiceless consonants raise the pitch. This effect has been noted in Chinese, Hottentot, and occurs even in English. House and Fairbanks conducted laboratory studies of the fundamental frequency of English monosyllables of the form CVC, where either both initial and final consonant were voiced, or both were voiceless.[23] Hence, contrastive pairs would be like bob : pop. Their findings show that the fundamental frequency of the voiced consonant syllables tended to be about 6Hz lower than the corresponding voiceless consonant syllable. The difference is not great, but one must remember that a syllable uttered in isolation has primary stress, and the latter is associated with high pitch in English. Unfortunately, they did not compare these curves with syllables where the initial and final consonants differ in voicing, so that the relative influence of the initial consonant to the final consonant on the pitch was not determined. With respect to this question, then, we shall discuss the tonal phenomenon of Jingpho Kachin.

Jingpho is a Tibeto-Burman language spoken in Burma, among other places. LaRaw Maran, in his thesis,[24] Tone in Burmese and Jingpho, states that the language has three tones, high, mid, and low. In closed syllables, where the final consonant is either a glottal stop or an obstruent, one finds only high or low toned syllables. With open syllables, and those closed by a nasal, one finds all three tones occurring. However, there are certain phonetic differences in the high and low toned open syllables as opposed to the mid toned syllables which do not appear to be related to the pitch. In syllables with a nasal final, the high and low toned syllables do not appear to have the strongly nasalized vowel which is found in the mid tone syllables. In syllables where the nucleus is a pure vowel, and where there is no final obstruent, the high and low toned syllables show a breathy vowel toward the end of the articulation which is not found in the mid tone syllables. Different degrees of nasalization and/or breathy vowels have never, to our knowledge, been correlated with any particular pitch. In syllables which are closed by an obstruent, there appears to be a contrast in voicing in the final obstruent. These differences of voicing are not easily detectable when the syllables are uttered in isolation, due to the fact that final obstruents in Jingpho are unreleased, but are quite apparent when the syllables occur with a clitic. If the clitic has no initial consonant, and if the preceding syllable is closed by a consonant, the final consonant becomes geminated. For example, one finds the following verbs: /yàg/ to be difficult, /ráp/ to cross, /tsam/ to wear. (´ indicates high tone, \` indicates low tone, and no mark is mid tone.) When these verbs occur with the post-sentential affirmative particle, /-ay/, which is a clitic, the resultant forms are: /yàg-gay/ it is difficult, /ráp-pay/ it (he) crossed, /tsam-may/ it is worn. In these latter examples, the final consonant of the verb is geminated and the voiced or voiceless quality of that final consonant is quite easily detected. No such gemination occurs when the word ends with the glottal stop. Maran cites similar examples to support the existence of an /h/ following the syllabic vowel in high toned open syllables. When a syllable ends in a simple vowel, and is high pitched, the existence of an /h/ is shown by the fact that any following clitic which has no initial consonant, for example /-ay/, shows an aspirated onset. Thus, he concludes that /h/ is also geminated.[25]

Given that there is reason to support the claim that there is a voicing distinction, at least phonetically, in final consonants, one then finds that there is a regular correspondence in obstruent closed syllables between the tone of the syllable, and the value of the feature [voice] of the final consonant. Specifically, syllables closed by a voiceless consonants are high toned, and those closed by voiced consonants are low toned. When studies were performed on the fundamental frequency of the vowels in these syllables, it was seen that the frequency ranges shown in these syllables corresponded to those shown in open syllables with high and low tone. They never corresponded to the frequency range acceptable for a mid toned syllable. We will discuss later the question of whether or not the voicing distinction in final consonants is an underlying distinction, or whether it is simply a phonetic consequence of the tone of the vowel. We will now discuss the acoustic findings of the pitch contours of the closed syllables.

In isolated utterances, the following was determined about the frequency curves of the different pitched syllables.[26] The high tone syllables had a steady frequency of about 160Hz. The syllables were short (duration of about 140 msec) when closed by an obstruent or glottal stop. When the syllable is closed by a real obstruent, i.e. not a glottal stop, the fundamental is generally somewhat lower. If the initial consonant is voiced, the fundamental frequency begins lower, at approximately 120Hz, and rises throughout the articulation of the vowel. With initial voiceless consonants, a frequency change is only observed if the syllable is closed by a real obstruent, not if the syllable ends in a glottal stop, nasal, or breathy vowel. The starting fundamental fundamental is, however, not as low as with voiced consonants.

The mid tone syllables show a level frequency curve of about 115Hz. We shall not discuss this except to remark that, as expected, the syllables with voiced initials tended to be slightly lower in frequency than those beginning with voiceless consonants.

The low tone closed syllables had an average duration of about 20 msec longer than the corresponding high tone syllables. The frequency curve is in general falling, with a sharper fall evident in syllables with an initial voiceless consonant. The lowest frequency generally attained is about 80 Hz, with the starting frequency in the neighborhood of 105 Hz. The syllables with voiceless initials generally had a starting frequency of 110 Hz to 120 Hz. One instance of the syllable /tà/ in low tone had a starting frequency of 155 Hz.

From the preceding, we find once again verification of the observation that voiced initial consonants induce a lower starting pitch. With open syllables (the Jingpho mid tone syllables), the effect was carried throughout the articulation of the vowel. In closed syllables, the effect is seen only on the starting frequency. We furthermore observe that with voiced initials, the frequency curve for high tone syllables is rising, and for low tone syllables with voiceless initials, it is falling, generally about 30 Hz or more. These contours are particularly evident since the pitch specifications for the vowel are identical to that of the following obstruent. Thus, for English, a primary stressed vowel must be high pitched irrespective of the nature of the consonants surrounding it. Consequently, the consonants cannot affect the pitch contour of the vowel to any great extent. On the other hand, in a language like Jingpho, where the pitch of the vowel agrees with the inherent pitch of the following consonant, the contours which are manifest can be considered as exaggerations of the natural transition from the pitch determined by the initial consonant to that determined by the final consonant. When the vowel articulation is very short, and the syllable is closed, with the initial and final consonants disagreeing in voicing, it is in fact difficult to conceive of a level frequency curve, since the transition from consonant to vowel and vice-versa must be rapid. We feel that the effect of consonants, particularly the final consonant, and stress probably account for the majority of the "contour" tones heard on short syllables.

One last word about short contour tones. Another source of contours may result from eliciting tones on single syllables in isolation. There is a tendency for speakers to utter isolated words with a peculiar "list" intonation, namely, either a rising intonation, or a falling intonation at the end of the word. With open syllables, especially in the Sino-Tibetan languages where open syllables have a long syllabic nucleus, the effects may not be particularly noticeable. However, with closed syllables, especially if the vowel is short, the upward or downward turn can become more noticeable because of its relatively longer duration in comparison to the total vowel duration. It is particularly striking that although there are reports of contour tones in syllables with short vowels, the contours never appear to be contrastive in the short syllables as they are in long syllables. Thus, although one will find languages which have a contrast between a mid to high rising tone and a high to mid falling tone on open syllables, one will not find such contrasting contours in closed syllables with a short vowel.

4.3.0 We shall now discuss briefly a question which has been of interest to linguists and propose two possible solutions. These solutions are not necessarily new, in fact, one has been proposed in slightly different forms by various sinologists, but we will present some evidence which seems to support that solution. How do languages develop lexical tones. Before answering this question, one must divide languages into two types: those whose phonological and morphological words are primarily polysyllabic, and those whose phonological and morphological words are primarily monosyllabic. In the former category are most of the tone languages of Africa and the Americas, and in the latter category are most of the tone languages of Asia. Let us look first at the first category and consider the example of Northern Tepehuan. There, one will recall, the position of the high tone is totally predictable and is a function of the length of the word. However, the part of the word which is considered for tone assignment includes only the stem and suffixes, not the prefixes. One can easily imagine that one may become partially unpredictable if some of the suffixes either disappear, or become shortened, but the tone configuration of the word maintained. On certain items which are trisyllabic, the speakers tend to place the high tone on the first syllable rather than on the second where it is predicted. An example is the verb /gagára/ to sell. In both the singular and the plural, one speaker pronounced the forms as /gágara/ and /gáágara/. Apparently, here there is some confusion between the plural and singular forms and the speaker has just chosen to place the high tone on the first vowel cluster. Thus, one way that a language could develop lexical tone is to have been at an earlier stage like Northern Tepehuan, with tone predictable, and then losing or changing part of the environment for predicting tone. Obviously, there must be other ways in which a polysyllabic language can develop tone. Let us turn now to the second category, the monosyllabic languages.

Several people have hypothesized that Chinese was not originally a tone language and/or that Archaic Chinese, the language of the Chou dynasty (around 1028 B.C.), had voiced final consonants. Among those who have hypothesized one or the other are Karlgren,[27] Simon,[28] Haudricourt,[29] and Forrest.[30] If early Chinese was not a tone language, then how did it develop into a tone language? It is generally accepted that by the time of Ancient Chinese, around 600 A.D., Chinese was a tone language, and that there were four tone contours, three found on the open syllables, and one unique to syllables closed by an obstruent. In addition, there was no voicing distinction in final consonants at this time. The three open syllable contours are said to be even, rising, and falling. Sinologists then say that a split occurred dividing these four contours into two series, an upper series for syllables beginning with voiceless consonants, and a lower series for those beginning with a voiced consonant. The remnants of this division are found in the Cantonese, Min, and Wu dialects. This latter development would coincide with our findings that voiced consonants depress the fundamental frequency of the Jingpho mid tone syllables uniformly. In particular, since at this point in the development of Chinese there was no longer a voicing distinction in the final consonant, the only consonant which would have had an effect on the contour would be the initial consonant. The lower frequency curves then became the distinguishing feature for some syllables when the initial voicing distinction was lost, or when segmental changes occurred in the initial consonant. The Thai languages experienced a further subdivision, this time governed by the feature of aspiration. Haudricourt has proposed a rather interesting theory as to the conditions under which a language will render these kinds of splits "phonemic" in his paper, "Bipartition et Tripartition des Systèmes des Tons dans Quelques Langues d'Extrême Orient". Beach reports that an analogous split occurred in Hottentot.[31] The tone splits have often been discussed in the sinological literature, and there is general agreement that the tone split in Chinese occurred as described above. On the earlier development of tone, however, very little has been said, primarily because of the innumerable disputes over any particular reconstruction of the segments.

Archaic Chinese has been reconstructed by both Simon and Karlgren as having voiced final consonants, although there is disagreement as to what these consonants were. Simon's system of voiced finals is actually more extensive than Karlgren's, so that the vast majority of reconstructed syllables are closed according to him. If we assume that these reconstructions are correct in positing final voiced consonants, and that proto-Chinese was not a tone language, then how might the tones noted by Ancient Chinese times have developed? Let us digress momentarily to examine Jingpho in more detail, since this is a modern Sino-Tibetan language which still shows a contrast between voiced and voiceless obstruents in final position.

We must first determine for Jingpho whether or not the voicing distinction in the final consonants is an underlying distinction. That is, are final consonants specified as both + and - [voice] in the lexicon, or is the quality of voicing in the final consonant a phonetic consequence of the tone of the preceding vowel? In the former case, it would then follow that the tone of the preceding vowel would be predictable from the value of the feature [voice] of the consonant. One might say that if the former is true, why specify any particular pitch value for the vowel at all? The reason for specifying a pitch level on the vowel comes from the fact that the pitch differences between the high, mid, and low toned syllables are greater than one expects if the contour results simply from the transition from consonant to vowel and then from vowel to consonant. Maran argues that the voicing distinction is an underlying one. If one considers only isolated syllables, then there is no reason to choose between the two alternatives. However, if one looks at various morphological processes, then the choice becomes evident. For example, in the morphology there are processes for the derivation of adjectives from verbs which involves an alternation of the final consonant and of the tone. Thus, one has the following pairs:

(3).	/ràb/ <u>to grade (or level)</u>	/ra/ <u>even, level</u>
(4).	/brùd/ <u>to shear</u>	/bru/ <u>shorn</u>
(5).	/sàd/ <u>to execute, do</u>	/sa/ <u>executable</u>
(6).	/gàd/ <u>to run</u>	/gan/ <u>running</u>
(7).	/sàd/ <u>to kill</u>	/san/ <u>lifeless</u>
(8).	/ràd/ <u>to space</u>	/ran/ <u>spaced</u>

This set of examples shows two kinds of alternations, in examples (3) - (5), the voiced consonant is deleted and the vowel lengthened, while in examples (6) - (8), the final obstruent becomes the homorganic nasal. In both sets, the derived adjectives have a mid tone while the verbs are low toned. There are other classes of alternations when adjectives are derived from verbs. All of them, however, involve some kind of a segmental change in the final consonant, and involve a concomitant tone change. Among these is the alternation of the glottal stop in low tone syllables with a glide in the adjective when the vowel is /o/.

Maran, in fact, claims that there are two sets of final laryngeals in Jingpho, /h/ and /ħ/, /?/ and /ꝯ/, the former member of each pair characteristically occurring with high tone syllables, and the latter member with low tone syllables. We will not discuss the reasons for conjecturing the two sets of laryngeals, but will refer the reader to Maran.[32] If one accepts Maran's statement, then the previously stated alternation is one between the "voiced" glottal stop and a glide, for Maran feels that /ꝯ/ and /ħ/ differ from /?/ and/h/ in an analogous manner to the difference between /b/ and /p/.

Another morphological process which involves an alternation of a final voiced stop with another segment is the process of deriving verbs from nouns. Here one finds that the low tone closed syllable alternates with the high tone breathy vowel syllable, that is, one closed by the"voiceless" /h/. Thus, one has:

(9). /bùd/ pile, heap /búh/ to multiply

(10). /sùd/ <u>wealth</u> /súh/ <u>to be (become) rich</u>

(11). /jìd/ <u>urine</u> /jíh/ <u>to urinate</u>

(12). /gùd/ <u>age</u> /gúh/ <u>to age</u>

Obviously, one could claim that the process goes in the opposite direction, but that is irrelevant. What is significant is the fact that not only does the tone change, but also the nature of the final segment. What do these morphological processes show? They show that if one considers tone to be underlying and the voiced quality of final consonants to be predictable, then one must have morphological rules which not only change the tone, but also specify the derived nature of the final consonant. On the other hand, if one says that tone is always predicted by the nature of the feature [voiced] of the final segment, then no morphological rule has to specify the change in tone, only the change in the final consonant. In particular, one will notice that the pitch is totally predictable from the nature of the final segment. Thus, if /d/ ~ /h/, then the vowel preceding /d/ is always low toned, and the vowel preceding /h/ is always high toned. However, given the fact that a vowel is low toned, one only knows that the post-vocalic segment is either a voiced obstruent or /h/ or /ʔ/.[33] Thus, in particular, in examples (6) - (8), from the knowledge that the vowel changes from low tone to mid tone, one cannot predict that the final segment will be the nasal homorganic to the stop. We therefore conclude that the voicing distinction for final consonants is an underlying distinction, and that tone is not lexical, at least for these syllables, but is entirely predictable from the value of the feature [voiced] of the post-vocalic segment. Maran, in fact, claims that the tone of all syllables is predictable from the nature of the post-vocalic segment in Jingpho, but we shall not go into the other cases here.

Let us return now to the problem of early Chinese. Aside from the fact that various linguists have reconstructed voiced final consonants, is there any reason to believe that early Chinese might be similar to Jingpho? If this should prove to be the case, then Haudricourt's hypothesis about the origin of Sino-Tibetan tone, for example, would gain support. Recall that the post-vocalic segment of high and low tone syllables in Jingpho is said never to be a simple vowel or nasal, but rather that there is always an obstruent or laryngeal. Recall also that when the initial consonant is voiced, the high tone syllables show a continuously rising fundamental frequency, and that all low tone syllables show a decreasing fundamental, with the fall being sharper when the initial is voiceless. What we wish to determine is whether or not there are any reasons to think that Archaic and/or pre-Archaic Chinese is similar to Jingpho and therefore, that the tone of Ancient Chinese may have developed as phonemic because of a loss of final segments or a change in the nature of the final segments.

We will consider Karlgren's reconstruction only, since his is the most widely accepted. Karlgren reconstructs the stops /b/, /d/, and /g/ as occurring in Archaic Chinese in final position along with their voiceless counterparts.[34] From the evidence of the rhyme dictionary written in 600 A.D., as given in the expanded version published in 1000 A.D., it is seen that the voiced series had apparently disappeared by Ancient Chinese times, that is, 600 A.D. However, since the time span between the two stages of Chinese is at least 1500 years, from 1000 B.C. (Archaic Chinese) to 600 A.D. (Ancient Chinese), this is not a particularly surprising development. Because of various considerations, however, Karlgren has concluded that many of the open syllables in Ancient Chinese were closed syllables in Archaic Chinese, and, in fact, closed by a voiced stop. Forrest,[35] in examining this reconstruction, has noticed several interesting points. The first is that the majority of the syllables closed by a voiced consonant appear in Ancient Chinese with the _ju sheng_ or falling tone. The second is the fact that there are fewer syllables in Karlgren's reconstruction of Archaic Chinese which are closed by a voiceless stop than by a voiced stop. We shall come back to this point later. The third observation concerns some semantic or possibly morphological relations between syllables closed with a voiceless obstruent and those closed with a voiced obstruent which are otherwise identical in their segmental composition (or nearly identical). Morphology is generally considered to be non-existent in modern Chinese although various people, among whom is Forrest,[35] have noted that there is a tendency to derive a noun from other grammatical categories by assigning the syllable a falling tone. However, Forrest points out that there are many exceptions to this tendency. From the material in Karlgren's syllabary, however, Forrest finds a large number of examples in which a morphological process seems to be indicated, in which the derived form, for him, is the one with a voiced final consonant. Downer has recently shown that the falling tone in the Han period (~200 B.C. to ~200 A.D.) vocabulary was associated with derived forms.[36] This would then support

Forrest's feeling that the voiced stop finals in syllables indicated that they were the derived form when they occurred in close relationship to a syllable without a voiced stop final. Forrest found three groups of words in which strong morphological-semantic correlations existed. The first group, the smallest, are those which were transcribed with the same character. Examples are the following (we will only cite a few):

(13). 805 /ʔâk/ bad

/ʔâg/ to treat as bad

(14). 496 /ţ'jwət/ to come out

/ţ'jwəd/ to cause to come out

(Forrest gives for this pair the glosses sortir-faire sortir)

(15). 921 /d'jək/ to eat

/d'jəg/ food

(16). 801 /dak/ to measure

/dag/ a unit of measure (length)

(17). 1120 /ʔjok/ to bind

/ʔjog/ bond, contract

(The numbers in the above examples correspond to the phonetic group in the Grammatica Serica.)

The correlations are of the type, nouns derived from verbs, causative verbs derived from adjectives or verbs, and reflexiveness. The second group of syllables in which a correlation was found were those in which the phonetic of the two characters was the same. The nature of the correlations is the same as for the first group. Examples are:

(18). 780 /d'âk/ <u>to inhabit</u>

/dag/ <u>to cause to enter one's house</u>

(19). 766 /klak/ <u>to arrive</u>

/glag/ <u>road</u>

(20). 279 /k'iat/ <u>to cut</u>

/k'iad/ <u>inscription, notch, groove</u>

(21). 695 /njəp/ <u>to enter</u>

/nwəb/ <u>interior</u>

The last group consists of words whose characters show no similarities. This is by far the largest group, numbering well over 130. Examples are:

(22). 1077 /d'jok/ <u>rapid</u>

1149 /d'jog/ <u>to accelerate</u>

(23). 925 /sjək/ <u>to cease, end</u>

977 /zjeg/ <u>end</u>

(24). 858 /liek/ <u>to calculate, enumerate</u>

878 /lieg/ <u>a number</u>

Obviously, the correlations are nowhere as neat as those of the Jingpho pairs, but that is to be expected, since these are reconstructed forms for which one can only obtain an approximate meaning. Nonetheless, there does appear to be a correlation, and one which is of a morphological nature. If we accept this, then it is clear that the morphological alternations, like those of Jingpho, principally affect the final consonant. If early Chinese does in fact resemble Jingpho, then it is conceivable that the tone system present by Ancient Chinese times arose from an earlier association of tone with final consonants and then the loss of the consonants. This is partially supported by the fact that those syllables reconstructed with a final voiced consonant appear for the most part with a falling tone in Ancient Chinese, just as the syllables with a voiced consonant final in Jingpho show a falling tone. What about the rising tone? Recall that Forrest noted that there were fewer syllables reconstructed for Archaic Chinese with voiceless finals than with voiced finals. On the other hand, every voiceless final that is reconstructed for Archaic Chinese appears in exactly the same form in Ancient Chinese, that is, dentals remain dentals, labials remain labials, and velars remain velars. In contrast, many of the modern dialects which still have closed syllables have changed the point of articulation in some forms, and even lost all remnants of a final stop in other forms. Thus, in Lungtu, Egerod finds that the combination

/-ək/ reconstructed by Karlgren for Ancient Chinese corresponds at times to /-ak/, /-it/ and /-et/.[37] In addition, he can find no traces of such combinations such as /-wak/. We presume this means that Ancient Chinese syllables with this nucleus have not only lost the final consonant, but have changed the vowel completely, and that they appear in Lungtu as open syllables.

It is therefore conceivable that Archaic Chinese had more syllables closed with voiceless final consonants than are listed, and that these consonants were lost by Ancient Chinese times. Recall that the time span between Archaic Chinese and Ancient Chinese is at least as great as that from Ancient Chinese to Lungtu. If there were more occurrences of voiceless finals, then the rising tone might have resulted from the loss of these consonants, especially since one has seen that in Jingpho, the syllables with voiced initials and voiceless finals show a steadily rising fundamental. The level tone would, of course, be associated with those syllables which had always been open, this category including syllables with nasal finals.

The "phonemicization" of tone, or making contrastive pitch a lexical feature of vowels due to loss of final segments is not unknown in languages. One sees this happening even in Jingpho. In the Gauri dialect, one finds breathy vowels in high toned syllables, and hence, can justifiably posit a post-vocalic /h/, from which the high pitch of the vowel results. In the Sirali dialect, however, there are no breathy vowels. Hence, for open syllables, one must consider the pitch of the vowel to be specified in the lexicon, for there is still a three way contrast in open syllables with respect to pitch.

CHAPTER IV: FOOTNOTES

1. Haudricourt, "L'Origine des Tons en Vietnamienne".

2. Chomsky and Halle, SPE, pg. 304

3. Zemlin, Speech and Hearing Science, pg. 188

4. Pronovost, "An Experimental Study of Methods for Determining Natural and Habitual Pitch".

5. see Stevens, Klatt, and Maran, "An Acoustic Study of Tone in Jingpho Kachin"

6. Zemlin, op. cit. pg. 191

7. My thanks to Eugene Ibe for the many hours of help with Igbo.

8. My thanks to LaRaw Maran who first suggested that the features [high tone] and [low tone] be correlated with the position of the glottis.

9. personal communication, LaRaw Maran.

10. Ladefoged, A Phonetic Study of West African Languages, pg. 5.

11. ibid. pg. 6

12. ibid. pg. 6

13. Pollack, "The Information of Elementary Auditory Displays", pg. 745.

14. Pollack, op. cit., figures 2 and 3, pg. 746.

15. personal communication, Ken Hale.

16. Ken Hale tells me that the long stressed vowels in Jemez do show a rising-falling contour, thus supporting the analysis given.

17. Longacre, "Five Phonemic Pitch Levels in Trique".

18. Chao, Mandarin Primer, pg. 26

19. Dreher and Lee, "Instrumental Investigations"

20. Öhman, "Word and Sentence Intonation: A Quantitative Model"

21. Smith, "Tone in Ewe".

22. Lieberman, Intonation, Perception, and Language, pg. 83 and pg 85.

23. House and Fairbanks, "The influence of consonant environment upon the secondary acoustical characterists of vowels".

24. Maran, Tone in Burmese and Jingpho, see especially chapter IV.

25. I must thank LaRaw Maran for the many hours he has spent with me discussing Jingpho and other tone problems. In particular, I must thank him for so patiently explaining the Tibeto-Burman problems.

26. Stevens, Klatt, and Maran, op. cit.

27. Karlgren, Compendium of Phonetics in Ancient and Archaic Chinese.

28. Simon, Zur Rekonstruktion der altchinesischen Endkonsonanten.

29. Haudricourt, op. cit.

30. Forrest, "Les occlusives finales en chinois archaïque."

31. Beach, The Phonetics of the Hottentot Language, pp. 247 ff.

32. Maran, "The finals in Tibeto-Burman"

33. Maran hypothesizes that with low tone syllables closed by a nasal, the syllable actually has the structure CVħN, where the existence of /ħ/ is supported by the fact that high and low toned syllables closed by a nasal show less (sometimes

34. Karlgren, op.cit.

35. Forrest, "Les occlusives finales en chinois archaïque", and The Chinese Language, pg. 125.

36. Downer, "Derivation by Tone Change in Classical Chinese".

37. Egerod, The Lungtu Dialect.

BIBLIOGRAPHY

Abramson, A.S. (1962), "The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments", International Journal of American Linguistics, 28.3.

Anderson, L. (1959), "Ticuna Vowels: with Special Regard to the System of Five Tonemes", Serie Linguistica Especial, no. 1, Publicações do Museu Nacional, Rio de Janeiro.

Bascom, B. (1960), "Tonomechanics of Northern Tepehuan", Phonetica, vol. 4, pp. 71-88.

Bascom, B. (1965), Proto-Tepiman, Ph.D. thesis, University of Washington.

Beach, D.M. (1924), "The Science of Tonetics and Its Application to Bantu Languages", in Bantu Studies, 2nd series, vol. 2, pp. 75-106.

Beach, D.M. (1938), The Phonetics of the Hottentot Language, Cambridge, W. Heffer and Sons, Ltd.

Bird, C.S. (1966a), Aspects of Bambara Syntax, Ph.D. thesis, University of California, Los Angeles.

Bird, C.S., (1966b), "Determination in Bambara", Journal of West African Languages, 3.1, pp. 5-12.

Bird, C.S. (1968), "Relative Clauses in Bambara", Journal of West African Languages, 5.1, pp. 35-47.

Bodman, N.C. (1955), Spoken Amoy Hokkien, vol. 1, Kuala Lumpur.

Brotzman, R. (1964), "Progress report on Mandarin tone study", Project on Linguistic Analysis, no. 8, pp. 1-35.

Browne, E.W., and McCawley, J. (1965), "Serbo-Croatian Accent", a translation of "Srpskohrvatski akcenat", Zbornik za filologiju i lingvistiku (Novi Sad), vol. 8, pp. 147-151.

Chao, Yuen-ren (1930), "A system of tone letters", Le Maître Phonetique, vol. 45, pp. 24-27.

Chao, Yuen-ren (1964), *Mandarin Primer*, Cambridge, Harvard University Press.

Cheng, Chin-chuan (1968), *Mandarin Phonology*, Ph.D. thesis, University of Illinois.

Chomsky, N., and Halle, M. (1968), *The Sound Pattern of English*, New York, Harper and Row.

Denes, P.B., and Pinson, E.N. (1963), *The Speech Chain* Bell Telephone Laboratories.

Doke, C.M. (1931), *A Comparative Study in Shona Phonetics*, Johannesburg, University of Witwatersrand Press.

Downer, G.B. (1959), "Derivation by Tone Change in Classical Chinese", *Bulletin of the School of Oriental and African Studies*, vol. 22, pp. 258-290.

Dreher, J., and Lee, Pao-ch'en (1966), "Instrumental Investigation of Single and Paired Mandarin Tonemes", Douglas Advanced Research Laboratory.

Egerod, S. (1956), *The Lungtu Dialect*, Copenhagen Ejnar Munksgaard, Ltd.

Forrest, R.A.D. (1960), "Les occlusives finales en chinois archaïque", *Bulletin de la Société de Linguistique de Paris*.

Forrest, R.A.D. (1965), *The Chinese Language*, London Faber and Faber.

Hanyu Fangyan Cihui, Peking, 1964.
(汉語方言詞汇)

Haudricourt, A.G. (1954), "De l'origine des tons en viêtnamien", *Journal Asiatique*, vol. 242, pp. 68-82.

Haudricourt, A.G. (1961), "Bipartition et tripartition des systèmes des tons dans quelques langues d'Extrême Orient", *Bulletin de la Société Linguistique de Paris*.

Hockett, C.F. (1947), "Peiping Phonology", in *Readings in Linguistics*, Chicago, University of Chicago Press.

House, A.S., and Fairbanks, G. (1953), "The influence of consonant environment upon the secondary acoustical characteristics of vowels", Journal of the Acoustical Society of America, vol. 25, pp. 105-113.

Jakobson, R., Fant, C.G.M., and Halle, M. (1963) Preliminaries to Speech Analysis, Cambridge, M.I.T. Press.

Jonggwo wentzyh gaeger woeiyuan huey, Peking, 1963. (中国文字改革委員會)

Karlgren, B. (1963), Compendium of Phonetics in Ancient and Archaic Chinese, Stockholm, The Museum of Far Eastern Antiquities.

Karlgren, B. (1964), Grammata Serica Recensa, Stockholm, The Museum of Far Eastern Antiquities.

Kiparsky, P. (1968), "How Abstract is Phonology?", mimeo, M.I.T.

Kratochvíl, P. (1964), "Disyllabic Stress Patterns in Peking Dialect", Archiv Orientální, no. 32, pp. 383-405.

Ladefoged, P. (1964), A Phonetic Study of West African Languages, West African Languages Monographs, Cambridge, The University Press.

Li, Kung-pu (1963), "Tone perception experiment with appended test materials", Project on Linguistic Analysis, no. 6.

Li, Yong-ming (1959), Phonology of Chaochow Dialect, Peking (李永明 潮州方言)

Lieberman, P. (1967), Intonation, Perception, and Language, Cambridge, M.I.T. Press.

Lo, Ch'ang-p'ei (1932), Phonetics and Phonology of the Amoy Dialect, Peking (羅常培 廈門音系)

Longacre, R.E. (1959), "Five Phonemic Pitch Levels in Trique", Acta Linguistica, vol. 7, pp. 62-81.

Maran, LaRaw (forthcoming), "The Finals in Tibeto-Burman".

Maran, LaRaw (1968), Tone in Burmese and Jingpho, Ph.D. thesis, University of Illinois.

Martin, S.E. (1957), "Problems of hierarchy and indeterminacy in Mandarin phonology", Bulletin of the Institute of History and Philology, Academia Sinica, vol. 29, pp. 209-229.

McCawley, J. (1964), "What is a Tone Language?", presented at the summer meeting of the LSA, 1964.

McCawley, J. (1965), The Accentual System of Standard Japanese, Ph.D. thesis, M.I.T.

Öhman, S.E.G. (1967), "Word and Sentence Intonation: A Quantitative Model", excerpt from STL-QPSR 2-3, pp. 20-57.

Pike, E., and Pankraz, L (1967), "Phonology and Morphotonemics of Ayutla Mixtec", International Journal of American Linguistics, 33.4, pp. 287-299.

Pike, K., Barrett, R.P., and Bascom, B. (1959), "Instrumental Collaboration on a Tepehuan (Uto-Aztecan) Pitch Problem", Phonetica, vol. 3, pp. 1-22.

Pike, K. (1964), Tone Languages, Ann Arbor, University of Michigan Press.

Pollack, I. (1952), "The Information of Elementary Auditory Displays", The Journal of the Acoustical Society of America, vol. 24, no. 6, pp. 745-749.

Pronovost, W., (1942), "An Experimental Study of Methods for Determining Natural and Habitual Pitch", Speech Monograph, Vol. 10, pp. 111-123.

Simon, W. (1927), Zur Rekonstruktion der altchinesischen Endkonsonanten, Berlin

Skinner, L. (1962), "Usila Chinantec Syllable Structure:, International Journal of American Linguistics, 28.4.1, pp. 251-255.

Smith, N. (1968), "Tone in Ewe", Quarterly Progress Report of the Research Laboratory of Electronics, Massachussetts Institute of Technology, no. 90.

Stevens, K., Klatt, D, and Maran, LaRaw (forthcoming), "An Acoustic Study of Tone in Jingpho Kachin".

Wang, W.S.-Y. (1967), "Phonological Features of Tone", International Journal of American Linguistics, 33.2, pp. 93-105.

Wang, W.S.-Y., and Li, Kung-pu (1967), "Tone 3 in Pekinese", Journal of Speech and Hearing Research, vol. 10, pp. 629-636.

Zemlin, W.R. (1968), Speech and Hearing Science, New Jersey, Prentice-Hall.

BIOGRAPHY

The author was born in 1943 in San Antonio, Texas. She attended the Massachusetts Institute of Technology from 1961 to 1969. She received her B.S. in mathematics in 1965, and her M.S., also in mathematics, in 1966.