This paper will address the very large topic of Chinese phonology. Obviously, this topic cannot be fully explored in a paper of this length. Rather than provide an exhaustive study, I shall attempt to give an introduction to Chinese phonology, and then note some particular problems in the study of Chinese phonology.

The first part of the paper will deal with some broad historical events that influenced Chinese linguistics. The second part will go into a detailed study of the phonemic inventory of Chinese, looking at the theories espoused by several different phonologists, and the advantages and disadvantages of each theorist's view. The third section will deal with tone in Chinese -- exploring different ways to define and describe tone and tone sandhi in Chinese. Finally, several recent phonological models will be presented and discussed in view of their ability to handle the data of Chinese.
History. (this outline of Chinese history comes from Pulleyblank, 1984)

Historically, in the development of the Chinese language, the capital has set the precedent for linguistic norms. The first Chinese empire, the Qin, came into being in 221 B.C. There is little linguistic evidence that dates from this period, of course. The first real clues we have date from the Han period. During that dynasty, the capital moved from Chang-an to Luoyang, and with the move, we see a change in the rhyming patterns. In the latter years of the Han dynasty, Luoyang was nearly completely destroyed in war, and was rebuilt by the succeeding dynasty, the Wei. From this renovated city, then, a local, colloquial speech emerged in the third century A.D. In 317, Northern China fell to non-Chinese invaders, and refugees from the fallen north moved south to Nanking, and founded the Eastern Jin dynasty. With this move, the educated language of the south became the standard Luoyang dialect, which was much different from the colloquial southern speech.

The split between north and south lasted some 270 years. In the beginning of the fifth century, the Northern Wei reunited the North of China, and reestablished the capital at Luoyang. Then, in 590, the north and south were unified, and the capital returned to Chang-an. So, because of the Chinese venture to the south, the linguistic standards of north and south, at least in educated speech, were not radically different.

In 601, the Qieyun was published. This was a rhyming dictionary, the first recorded phonology of Chinese, and combined the similar elements of northern and southern educated speech. The publication of this dictionary marks what is referred to now as the Early Middle Chinese (EMC) period.

The local dialects of Chang-an influenced the standards of EMC, and by the seventh century, a new standard emerged, which is now known as Late Middle Chinese (LMC).

During the great T'ang dynasty, LMC gained a lot of respect, and began replacing more dialects as the Chinese standard language. This recognition of a national language of China took place not only within China itself, but also in neighboring countries -- Japan, Korea, and Vietnam.
The capital was moved to Kaifeng during the Northern Song, then to Hangzhou during the Southern Song, but the established northern standard remained.

The next major changes took place during the Yuan dynasty, when the capital moved to Peking. Under the rule of non-Chinese, the Chinese literary tradition began to dissolve, and the local dialects around Peking influenced the standard in many ways.

With the Ming restoration, the capital moved south to Nanking, and the language remained relatively stable, as the capital returned to Peking just 35 years later.

Since then, the capital in Peking has influence the standard language. During the Qing dynasty, this language was known as *guanhuān*, or Mandarin. When the Qing fell and the nationalists came into power, the language was called *guoyu*, or 'national language.' And, when the communists chased the nationalists to Taiwan, Chinese was referred to as *pǔtōnghuā*, or 'common language.'

Some changes that took place over this long, broad period of time.

In Old Chinese, [r] could occur as the second element in syllable-initial consonant clusters. So, clusters like [pr-], [tr-], [kr-], and so forth, were possible. By EMC, the [r] did not occur after labials or velars, but left its traces behind in retroflexing some vowels. In Middle Chinese, the unisegmental affricates [tr-], [tr'-], [dr-], [ts-], [ts'], and [dz-] had emerged. (the ['] in these examples indicates aspiration). Similarly, the initial cluster [kw-] had simplified into the labialized velar consonant (which I will also represent as [kw-] for typographical convenience, the distinction being made is that the first [kw-] is a consonant cluster, the second is one segment).

In the Qing dynasty, a rhyme phonology was developed which classified vowels in the following ways.
kaikou 'open mouth'
hekou 'closed mouth'
qichi 'level teeth'
cukou 'pursed mouth'

(Pulleyblank, 1984, p. 15)
These terms are similar to the features [front] and [labial] that are currently used to describe vowels. 'Open mouth' described those sounds that were made without lip rounding, and without showing the teeth, namely [-front], [-labial] vowels. 'Closed mouth' were those sounds that needed lip-rounding, but not pursing -- [-front], [@labial] sounds. 'Level teeth' meant that the teeth showed when the sound was being produced -- [@front], [-labial]. And 'pursed mouth' was the description for front rounded sounds, sounds that really need tense lips.

In addition to this classification, of course, we must add the features [high] and [low] to fully describe the Chinese vowels.

Further, the consonants of Late Middle Chinese were classified by a five-fold system, as follows.

yayin 'back tooth sounds'
sheyin 'tongue sounds'
chunyin 'lip sounds'
chiyin 'front tooth sounds'
houyin 'throat sounds'

These classifications, which referred mostly to place of articulation, were arranged in the following table.
In this representation, the 'back tooth sounds' refer to velars; 'lips sounds' are labials -- the first row of labials being 'heavy,' or bilabial, the second row being labiodental fricatives. 'Tongue sounds' are a bit more complicated. The first row are alveolar stops; the second row of the 'tongue' along with the second row of the 'front tooth sounds', according to Pulleyblank, are all sounds that come into modern Mandarin as retroflex sibilants, and his theory is that by LMC, palatals and retroflex consonants had merged. The first row of 'front tooth sounds' are easily recognized as alveolar affricates and fricatives, and the last group, the 'throat sounds,' probably consisted of laryngeals and uvulars.

Additionally, these columns were divided by phonation types. The exact meanings of the terms 'clear' and 'muddy' has been somewhat debated. Pulleyblank argues for a distinction based on voicing and aspiration, rather than on voicing alone, as had been previously claimed. "'Clear' did mean voiceless, but 'muddy' must be interpreted as 'having voiced aspiration' or 'murmured' rather than simply 'voiced.'" (Pulleyblank, 1984, p. 67). This distinction is important to separate voiced aspirates from voiced sonorants, as sonorants did not aspirate.

<table>
<thead>
<tr>
<th></th>
<th>clear</th>
<th>second</th>
<th>muddy</th>
<th>clear-muddy</th>
<th>clear</th>
<th>muddy</th>
</tr>
</thead>
<tbody>
<tr>
<td>yayin</td>
<td>k</td>
<td>k'</td>
<td>kh</td>
<td>η</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sheyn</td>
<td>t</td>
<td>t'</td>
<td>th</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>tr</td>
<td>tr'</td>
<td>trh</td>
<td>nr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chunyn</td>
<td>p</td>
<td>p'</td>
<td>ph</td>
<td>m</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>f</td>
<td>f</td>
<td>fh</td>
<td>u</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chiyin</td>
<td>ts</td>
<td>ts'</td>
<td>tsh</td>
<td>s</td>
<td>sh</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ts</td>
<td>ts'</td>
<td>(t)sh</td>
<td>s</td>
<td>sh</td>
<td></td>
</tr>
<tr>
<td>houyn</td>
<td>?</td>
<td>x</td>
<td>xh</td>
<td>(H)</td>
<td>l</td>
<td>r</td>
</tr>
</tbody>
</table>

(from Pulleyblank, 1984, p. 63)
Phonologies of Chinese -- the vowels.

There are many theories that differ in minor ways about the phonology of Chinese vowels. The traditional view is that Chinese has six underlying phonemes, from which all surface forms are derived. Some alternate theories propose five phonemes, or, in the extreme case, Pulleyblank's 'seven or nothing' phoneme theory, which we shall discuss later.

The easiest way to test various theories is to first chart all the Chinese finals, which include the syllabic unit and coda, and then test the theories against this framework.

Following is Pulleyblank's ordering of the Chinese finals.

<table>
<thead>
<tr>
<th>neutral</th>
<th>front</th>
<th>labial</th>
<th>front-labial</th>
</tr>
</thead>
<tbody>
<tr>
<td>r, z</td>
<td>i</td>
<td>u</td>
<td>ü</td>
</tr>
<tr>
<td>v, ð</td>
<td>ie</td>
<td>u₂</td>
<td>üe</td>
</tr>
<tr>
<td>A</td>
<td>iA</td>
<td>uA</td>
<td></td>
</tr>
<tr>
<td>ei</td>
<td>iei</td>
<td>uei</td>
<td></td>
</tr>
<tr>
<td>ai</td>
<td>iai</td>
<td>uai</td>
<td></td>
</tr>
<tr>
<td>ou</td>
<td>iou</td>
<td></td>
<td></td>
</tr>
<tr>
<td>αu</td>
<td>iαu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in</td>
<td></td>
<td>ün</td>
<td></td>
</tr>
<tr>
<td>ðη</td>
<td>iðη</td>
<td>uðη</td>
<td>üan</td>
</tr>
<tr>
<td>an</td>
<td></td>
<td>uan</td>
<td>üan</td>
</tr>
<tr>
<td>ηη</td>
<td>Ωη</td>
<td></td>
<td>iΩη</td>
</tr>
<tr>
<td>Κη</td>
<td>iκη</td>
<td>uΛη</td>
<td></td>
</tr>
<tr>
<td>αη</td>
<td>iαη</td>
<td>uαη</td>
<td></td>
</tr>
</tbody>
</table>

(Pulleyblank, 1984, p. 45)

The Ω in Pulleyblank's system, which he writes as Q, is an allophone of /u/, that surfaces only before the final η. Pulleyblank does not describe this sound.

The traditional Pinyin approach to Chinese vowels posits six underlying vowels -- three high vowels and three non-high vowels.
These vowels appear in various combinations to account for all the finals. For example, /i/ shows up in /ia/, /ie/, /iao/, /iou/, and so forth. Pinyin is not a true phonology of Chinese, however. It is instead the most universal type of romanization used in teaching Chinese. Therefore, its phonemic inventory does not attempt to rigorously describe all the sounds of Chinese, but it provides guidelines for the way in which vowels are pronounced in various environments.

Charles Hockett (1947) was probably the first to do a comprehensive review of the phonemes of Chinese. His paper includes a list of initials, a list of finals, and the rules required to join the two. His system is based on a sequence of features so, rather than positing underlying morphemes and deriving their surface structures, he describes the sounds as they occur. Thus, his 'vowel inventory' consists of the 'features' /i, u, r, e, a/ where /e/ and /a/ denote vowels with mid and low tongue height respectively; /i/, /u/, and /r/ denote semivowels with high front tongue position, lip rounding, and retroflexion, respectively. By this system, a combination of /ue/ would be the simultaneous articulation of lip rounding and mid tongue height, producing the [o] sound. (Hockett, 1947, p. 258).

Robert Cheng (1966) wanted to return to the concept of vowel phonemes, and simplify the Pinyin system. For him, it was necessary to have only one underlying mid vowel phoneme -- all the surface forms could be seen as allophonic variations of that form. His vowel chart looks like this:

```
  high   i   ü   u
  mid    e
  low    a
```

To begin with, R. Cheng claims that the two apical vowels [r] and [z] are not syllabic consonants, but are allophones of the phoneme [i]. According to his theory, [z] surfaces after dental affricates, [r] (retroflexed) surfaces after retroflexed affricates, and
[i] surfaces elsewhere.

The justification for having only one low vowel is not terribly complicated -- [a] surfaces before or after front [i] or [n]; and the back variant [α] surfaces before or after [u] or [n].

Then, R. Cheng justifies his claim that there need be only one mid vowel. /e/, he claims, surfaces as [e, e, ə, ʊ, ɔ, null]. His rules to justify this are understandably complicated.

1. /e/ ⇒ [e] / u_i
2. /e/ ⇒ [ə] / ʊ_i, i_
3. /e/ ⇒ [ʊ] / _u, i_u,
4. /e/ ⇒ [ʔ] / u_
5. /e/ ⇒ [y] / velar _ #

He sees these rules as assimilation rules. /e/ surfaces as [e] or [ə] before or after high non-back vowels; as [ʊ] or [ ə] before or after high back vowels, as [y] finally after a velar initial, and as [ʔ] in all other locations. Looking back at our data of Chinese finals, we can see that R. Cheng’s rules indeed account for all the data. But is his analysis the best possible? Why in rule 3, for example, does [ʊ] take precedence over [i] in determining the backness of the underlying /e/? R. Cheng would respond that regressive assimilation is predominant over progressive assimilation, but this seems like a tenuous justification. Also, there seems to be no reason for [e] surfacing as opposed to [ə], or for [ʊ] as opposed to [ʔ].

C.C. Cheng goes into more detail about the exact rules required to change his underlying forms into surface forms. Underlyingly, he proposes four high vowels; [i], [ʊ], [ɨ], [u]. He agrees with R. Cheng that [z] and [r] are not syllabic consonants, but are allophones of each other. For some reason, however, he doesn't see them as allophones of /i/; instead, he needs another high phoneme [ɨ], to account for the two variants. [ɨ], he claims, surfaces as [z] after strident dentals [j,q,x], and as [r], after
C.C. Cheng continues with two underlying nonhigh vowels, the low back /œ/ and the mid back unrounded /y/. These two forms are subject to the following series of rules.

The Backness Rule. "Within the final, a nonhigh vowel is assimilated in backness to the immediately neighboring segment, regressive assimilation being dominant over progressive assimilation." (C.C. Cheng, 1971, p. 14).

With this rule, C.C. Cheng accounts for some of the variation in nonhigh vowels. For example, the two low vowels are now definite allophones, with the same distribution as R. Cheng posited; [a] occurring before or after front [i] or [ɛ], and [œ] occurring before or after the back [u] or [ŋ]. C.C. Cheng has a bit of a problem, however, accounting for the variation in the [iæn] final. This final goes through Backness to emerge as [ian], but, at least for some people this final is phonetically [iɛn], although it continues to rhyme with the -an finals. So C.C. Cheng proposes an optional rule: [a] ⇒ [ɛ] / i _n, that seems again to be little more than an ad hoc justification for his theory.

This takes care of the low vowels and of the sole occurrence of [ɛ], so C.C. Cheng has only to account for the variation among the remaining mid vowels. With the backness rule, he can separate [e] from [o], and all that he has left is to account for the schwa. This he does with the following two rules.

The Mid Vowel Laxing Rule. "Within the final, a mid vowel before a consonant ending becomes a schwa." (C.C. Cheng, 1973, p. 18).

The Schwa Deletion Rule. "Within the final, a schwa between a front medial and a nasal ending or between a back medial which is preceded by an initial consonant and a back nasal ending is deleted." (p. 18).

With all of these rules, C.C. Cheng justifies his version of the six vowel phonemic system.
The problem with these rules, as I see it, is that, while they do account for the data, they don't necessarily justify themselves. Why, for example, should a mid vowel become a schwa before a consonant ending? And what motivation does a schwa before a nasal ending have for deleting? Unstressed vowels often become schwas, and unstressed schwas often delete, but these are not the justifications that C.C. Cheng calls upon to support his rules.

Pulleyblank's theory, as we stated before, is the 'seven vowel phonemes or none' theory. In his earlier writings, he critiques the lumping together of many so-called allophones to reduce the phonemic inventory. In his 1983 article, however, he takes phonemic reduction of vowels to its maximal point, postulating that Chinese can, in fact, be treated as a 'vowelless' language -- underlyingly having only glides and consonants, which surface as vowels. We shall look at this theory in detail, but first let us consider his more traditional approach.

Glides are a key factor in Pulleyblank's analyses, but he can't claim to have been the first to fit them into Chinese phonology. Quite ahead of Pulleyblank, Hartman (1944) arrived at some similar solutions to the phonology of Mandarin vowels. Hartman posits three underlying syllabics, along with the semivowels /J/, /W/, and /R/. The semivowels, he claims, occur in the medial position, and interact with the syllabic unit. Thus, he derives [i] from the glide /J/ plus /i/; [u] from /W/, and so forth. This paves the way for Pulleyblank's all-encompassing analysis of vowel segments in which the parts of the segment are attached to the same node, and pronounced simultaneously.

Pulleyblank expands the glide theory quite a bit. Glides and vowels, he claims, have always been linked phonemically, a linkage which he finds erroneous. The finals -η and -ωη, for example, are in complementary distribution -- -η occurring only after a consonant, and -ωη occurring with the zero initial. His explanation is that the 'u' in -ωη is actually a W glide, and functions here as the initial. Additionally, he postulates that the high vowels /i/, /u/, /ü/, and the low vowel /a/ can be treated as syllabic forms of glides. Making this distinction between glides and vowels universal, he proscribes the following set of rules.
(a) i and u are replaced by J and W as the final elements in diphthongs.
(b) i is replaced by J before all vowels except e in -ie and -ien.
(c) u is replaced by W before other vowels in all finals except u_
(d) ü is a glide in -üan.

In order to solve the problem of the phonemic values of the mid vowels and how to relate them with their rhyming patterns, Pulleyblank suggests that there is a final glide in the mid and low vowel finals. This glide is a low unrounded glide, which we will write as /ã/. (Pulleyblank writes it as /ã/). Thus, the mid vowels are analyzed as -åa, -Jå, -Wå, -üå.

Pulleyblank then moves on to analyze the low vowels. C.C. Cheng's system of fronting is inadequate, he claims. Changing /œ/ to [a] by a fronting rule simply does not work, as [a] is not /œ/\front.

So, Pulleyblank goes on to put his glide into the UR, positing /aå/ for [œ], /Jå/ for [ie], and /Jån/ for [ien], thereby relating -ien to its rhyme-mate -Jan, and avoiding the ad hoc rule that C.C. Cheng had to insert.

Further, the syllabic [z] and [r], Pulleyblank argues, are just that. Some people would categorize them as vowels; he claims it more economical to posit them as "syllabic extensions of the preceding initials." (Pulleyblank, 1986, p. 51).

There is historical justification for this, too. Historically, the syllabic [r] arose because of the deletion of [i] after retroflexed sibilants. With the deletion of the vowel, syllabicity was preserved in the preceding consonant.

So, with all these modifications of traditional vowel theory, Pulleyblank comes up with the following table of Chinese finals.
In this table, the vowel inventory seems to make more sense. The high vowels remain as before -- /i/, /u/, and /ü/; the /\theta/ and /\zeta/ are seen as syllabic consonants rather than high vowels. There is only one mid vowel, /\partial/, and one low vowel, /a/, but rather than compiling a set of eclectic rules to account for surface variations, Pulleyblank combines these two phonemes in various places with various glides to reach his surface destinations.

[e] and [o] are in complementary distribution with [\partial], and, Pulleyblank claims, "It seems legitimate to derive them from an underlying /\partial/ by a simple rule of assimilation to the following glides, -J and -W..." (Pulleyblank, 1984, p. 46). He goes on to look at the other mid vowels -- y, e, and ə. These vowels cannot be seen as allophones of /\partial/, he claims, because, with the retroflex suffix, -\partial remains distinct from -y. Additionally, the e in ien seems to come from /a/, as the ending ien rhymes with -an, -uan, and -üan. This is the problem that vexed C.C. Cheng, and forced him to come up with his a -\rightarrow e / i_\_n rule. Pulleyblank treats the problem differently, by simply not positing these mid vowels as allophones of /\partial/.

Later (his 1983 article was actually written before his 1984 book) Pulleyblank goes on
with his discussion about the possibilities for an underlyingly vowelless Chinese.

Following are the main points of his argument.

Every syllable in Mandarin, Pulleyblank claims, begins with one of his initials. The so-called 'zero initial' is, in fact the laryngeal glide /H/, which is pronounced by most speakers as a "frictionless velar or uvular continuant" (Y.R.Chao, quoted in Pulleyblank, 1984, p. 572).

[i] and [u] arise when the glides /J/ and /W/ associate with the V node of the syllable template; in other words, these high vowels are nothing but fully syllabified glides. Further, [W] is a labial initial, like /p/, /ɻ/, /m/.

All labial initials have a w "which never surfaces as a w glide but which can vocalize as u under certain circumstances." (Pulleyblank, 1983, p. 576). This labialization has historically made itself felt in such processes as dissimilating final labial after labial initials.

Further, the schwa that most people would claim is underlying in finals such as /iən/, /ən/, /wən/, is actually not present in the underlying representation, but is inserted between consonants as required. Pulleyblank does not go into detail about exactly when this happens.

To further evaluate his glide-oriented theory, Pulleyblank discusses the arguments made by Clements and Keyser about the question of French liaison. Clements and Keyser claim there is an underlying final consonant that is not attached to a syllable node, but is instead free-floating. This consonant can attach to the following syllable of a vowel-initial word and cause liaison, but it remains unattached before a consonant. Further, in 'h-aspiré' words, there is already an empty C-node, which causes the floating consonant to remain unattached. (Pulleyblank, 1984, p. 588). He gives the following examples.
In the second example, the 't' is the floating consonant, and links with the vowel of the following syllable, causing liaison. In the first and third examples, the initial C node blocks the liaison.

"This is surely very curious" Pulleyblank exclaims (p. 588). An empty C node (for reasons he doesn't feel the need to expound) should encourage, rather than block liaison. His solution is to claim again that initial aspiration is not merely an empty node, but is instead the laryngeal glide [H]. Vowel initial words, then, can be seen as having the empty C node, which is occupied by the [H] when the word occurs by itself, but is otherwise filled by the preceding consonant, thus leading to liaison.

If this is true, and both 'h-aspiré' and vowel initial words have their first node filled by the laryngeal glide [H], how does the phonology distinguish between the two, and make the liaison in one case, and fail to make it in the other? Pulleyblank claims that vowel initial words do not have the glide as part of their underlying representation; that, in fact, the first C node is empty, and the [H] only occurs to provide a "smooth, voiced transition" when the word occurs in isolation. In other words, the [H] glide is the default value for vowel-initial words. 'H-aspiré' words, on the other hand, have this [H] underlyingly, and therefore, the node cannot be filled by anything else.

Pulleyblank goes on to postulate the existence of empty V nodes, for, as he says, if C nodes can be empty, why not V nodes as well? This would account for the so-called 'mute e' in French. For example, the word *cela* is sometimes pronounced /sələ/, and sometimes /sla/, but is distinctive from the word *slav*, which cannot occur with the inserted /ə/. The difference here, Pulleyblank claims, is that *cela* has an underlying empty V node, which can be filled by the schwa, should occasion arise. The word *slav*, on the other hand, has no empty V node, so schwa insertion would be impossible. Similarly, in a word such as *l'enfant*, the definite article has an empty V node which
combines with the empty C node of *enfant* to form one syllable. However, a word like *heros* has its initial C node filled by the glide [H], and therefore cannot accept the empty V node of the definite article, so a schwa is inserted, giving /iðHer soar/ (Pulleyblank, 1984, pp. 590-592).

Complex segments in Chinese, such as labialized consonants, occur due to a simultaneous pronunciation of consonant and glide. For example, labialized labials, such as /pw/ or /bw/ are not pronounced linearly, but are simultaneous articulations. Taking this for granted, Pulleyblank then goes on to postulate the existence of simultaneously articulated sequences of glides and vowels, assuming that vowels differ from glides only in that they are dominated by a V rather than a C node. This accounts for the front rounded glide [u], which is the simultaneous articulation of /J/ and /W/, dominated by a C node, and its syllabic counterpart [Y], which is the same combination, only dominated by a V node, so that it looks like:

\[
\begin{array}{c}
V \\
\backslash \\
_i u
\end{array}
\]

(Pulleyblank, 1984, p. 597)

Is there justification for claiming that /4/ and /Y/ are combinations of /Jw/ and /[iu]/? Pulleyblank thinks so, showing that /4/ and /Y/ are in relatively complementary distribution with /J/ and /i/. Both incur palatalization of velars and alveolars that precede them. Also, if /4/ is in fact a segmental cluster of /Jw/, it can easily be explained as surfacing as either /Y/ (in -Yn and -Y) or as /Yu/ (in -Jun), when it appears under a C node. (Pulleyblank, 1984, p. 598).

There are, by Pulleyblank’s table, three kinds of finals in Mandarin -- those with no final consonant, which end in syllabic [z] or [r] or one of the high vowels; those which end in one of three possible final consonants and may have a schwa (or one of its allophones e or o), or a high vowel; and those which have a low /a/ vowel of some kind. (Pulleyblank, 1984, p. 600).

In the first group, it seems obvious to Pulleyblank, the syllabic [z] and [r] are cases of the spreading of the features of the initial consonant onto an empty syllable nucleus. There is no reason, given this analysis, to posit an underlying vowel here. Taking this
analysis a step further, he goes on to postulate that all syllables in Chinese take the CV form C(C)VC, and that in these cases, both the final C and the V node are left empty. If we then extend this template to the labialized initials discussed above, we get the following results.

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CVC</td>
<td>CVC</td>
<td>CCVC</td>
<td>CCVC</td>
<td>CCVC</td>
<td>CCVC</td>
</tr>
</tbody>
</table>
| ts  | s   | [pJ] | [pW] | kW  | [IJJ]
| [tsr] | [sr] | [pi] | [pu] | [ku] | [Ilu] |

(Pulleyblank, 1983, p. 601)

In the first two examples, the [ts] (which is a single segment) and the [s] don't have the optional glide, so the syllabic [r] is the syllabic extension of the initial. In the last four examples, the glides syllabify into their corresponding vowels.

Pulleyblank goes on to describe the history of the syllabic /r/. In the case of the word 'two' -- [r]; in the 8th century (LMC), it was [ri], and by Early Mandarin (14th century), had gone through the rule that dropped [i] after retroflexed initials, becoming [rr]. The word 'day' -- [rr], on the other hand, came from the root /rit/ in LMC, and didn't lose its final stop until EM, after the rule deleting /i/ after retroflexes had occurred. So, it remained in EM as /ri/. Since then, [rr] has become [r], and [ri] has gone to [rr].


The second set of finals has both initial and final C slots filled, containing both an initial glide and a final consonant. Additionally, the V node is filled by only a schwa or one of the high vowels found in the first set. The same distribution of /i/ and /u/ alternating with /i/ and /u/ plus shwa occurs as did in the first set and, as discussed before, Pulleyblank prefers here to posit an underlying zero, rather than a schwa. So, the V slot is not filled in the underlying representation, but gets filled either by having the vocalic features of the first glide spread (as they did in the first set), or by schwa insertion. The way in which the glide will vocalize is almost predictable by the frontness of the final consonant. /I/ becomes [i] before the neutral glide /å/ or front /n/, but not before /W/. /I/ becomes [u] before /å/ or /I/, but not before /n/ or /I/. We run into a bit of a problem, of course, when contrasting /I/ and /W/ combine as /U/, but this is
a small problem.

Additionally, leftward spreading from the final glide à accounts for the actual pronunciation of the finals [ià] and [uà]. The medial in these finals is, of course derived from the corresponding glide, so the association of /J/ and /W/ with the glide /à/ (under the same V node) results in the low front vowel [e] and the low back rounded [ɔ], respectively.

The final set of finals has the low vowel /a/ between the initial and final consonants. This /a/ is assumed somehow to be a vocalization of the V node -- needed here to preserve syllabicity. This seems to be another weak node of Pulleyblank's absolute argument -- in a sense /a/ must exist as an underlying vowel, even if its value is as a default, to be used if the V node needs to vocalize. But, Pulleyblank argues, the /a/, which he says is underlyingly a nonsyllabic glide, is associated first with an empty C node, and then with the V node to preserve syllabicity. In any case, the /a/ surfaces as [ɔ] before [n] or [W], a case of backness assimilation.

So, his conclusions are as follows. First, there exists a voiced laryngeal glide H, which surfaces as a consonant in syllable initial or final positions. Second, vowels and glides can be combined as complex segments, making intermediary vowels out of the primaries -- i,u,a,ə. And last, Chinese has a C(C)VC syllable structure, differing from the 'universal unmarked form' of CV proposed by Cairns and Feinstein (Pulleyblank, 1983, p. 614).

Deborah Davison, in her study of dialects and secret languages (1983), supports Pulleyblank's claim for a vowel insertion rule that accounts for the schwa insertion and the existence of the apical vowels.

Where Hockett proposes vocalization of the semivowels, and C.C. Cheng posits an underlying mid vowel ə, Pulleyblank argues that neither are present, but both are epenthetic, and come into play to restore syllabicity when needed. (This is a bit different from Pulleyblank's glide-to-vowel assimilation process, which is similar itself to Hockett's semivowel vocalization....) So, her interpretation of Pulleyblank continues,
the Mandarin syllable either has underlying /a/ or no vowel at all. (This seems to refute Pulleyblank's 'vowelless' Chinese, but at the same time provides a more adequate justification for the way in which low vowels arise...) Glides vocalize only if they are pre-vocalic and if the following consonant is of the same frontness. THEN, a dissimilation rule deletes [u] after a labial initial, and this is where the epenthetic schwa must be inserted to preserve syllabicity.
Phonology -- the initial consonants

Chinese initial consonants are far easier to deal with than the vowels, and there is consequently far less disagreement among theorists about the nature of initial consonants.

The following initial consonants are generally accepted for Chinese phonology.

- velar: g, k, h
- palatal: j, q, x
- retroflex: zh, ch, sh, r
- alveolar sibilants: z, c, s
- dental: d, t, n, l
- labial: b, p, f, m

There is also a 'zero' initial.

(C.C. Cheng, 1973, pp. 32-33)

The one point of contention is the exact nature of the initial palatal series. Pinyin's listing of the possible syllables in Chinese has the palatales in complementary distribution with the velars -- palatales occur before [i] and velars occur in other places. This is in fact the case; however, palatales are also in complementary distribution with retroflexes and alveolars, all of which do not occur before the high front [i]. Pulleyblank gives the following historical account which sheds some light onto, but doesn't solve the issue of, the palatal series. Palatales arose, he claims, from the palatalization of velars, then alveolars, in front of high front vowels. This is, of course, assimilation. Around the same time, high front vowels assimilated to the retroflexes that they followed, becoming the 'apical' vowels discussed above. (Pulleyblank, 1983, pp. 575-576). Pulleyblank supports his claim with some data from the ancient language of the Peking opera, in which [i] and [ü] do occur after retroflexes. (Pulleyblank, 1984, pp. 65-66) Since Pulleyblank does not put as high a price on extra phonemes as other theorists, he has no problem treating the palatales as separate phonemes.
Pulleyblank, of course, must revise the traditional analysis of initial consonants to fit into his theory. The zero initial, for example, as discussed above, is actually a voiced laryngeal, which surfaces as a velar or uvular approximate before non-high vowels. This approximate assimilates /a/ preceding /n/ to the velar /ŋ/.

Additionally, the 'h' is in actuality a uvular sound, which has a voiced counterpart, which we will write 'H'.

Glides, Pulleyblank argues, should also be included in the list of initials. Glides can occur finally or medially in finals, such as -aj, -eJ, -aW, -oW, -Jan, -Jow, -Wən, -WaJ. When these finals take a zero initial, there is a palatal or labial onset, J or W. Initial W is often realized as the approximate u.

So, Pulleyblank's revised list of initials is more symmetrical.

| velar/uvular | g | k | h | H |
| palatal       | j | q | x | J |
| retroflex     | zh| ch| sh| r |
| alveolar      | z | c | s |   |
| dental        | d | t | l | n |
| labial        | b | p | W(u) | m |
Tone.

Historically, Hombert et. al. (1979) argue, tones have arisen for a couple of different reasons. One is that a speaker's speech would be distorted; a listener, unable to determine the speaker's intention, repeats, exaggerates distortion. Over time, this exaggeration is repeated, until it conveys the meaning in question. Another reason they give is that tones evolve because of a loss of voicing on pre-vocalic obstruents. Vowels after voiced obstruents are articulated lower than those after unvoiced obstruents, so that, when the voicing distinction is lost, the height of the vowel articulation is exaggerated to make the distinction. Additionally, they argue, tone can originate from the influence of postvocalic glottals, but never from post vocalic non-glottals, or from distinctive vowel height (Hombert et al, 1979, p. 37).

McCawley (1978) has a brief and fun overview of Chinese tonality. Chinese has four tones, first (high), second (rising), third (either falling/rising or level and low), and fourth (falling). Additionally, there is a 'neutral tone' assigned to unstressed syllables, which surfaces as high if the preceding tone ends high and the following tone starts high, or if preceded by a third tone and followed by third or fourth tone. In other locations, it is realized as low. Then, there are two sandhi rules, one which changes third tone to second tone before third tone, and one which changes second tone to first tone when preceded by a first or second tone and followed by a stressed syllable in rapid speech.

McCawley then goes on to sum up the characteristics of pitch in Mandarin with three rules. First, tones are assigned to morphemes in the lexicon. Second, rules that affect pitch are assimilation and dissimilation rules that are effected by neighboring segments. Third, there is a distinction between stressed and unstressed syllables, and the tone on unstressed syllables is predictable from the neighboring tones. (McCawley, pp. 119-121).

As we shall see below, the question of tone in Chinese can be much more complicated than this.

The number and shape of Mandarin tones are seldom disputed. Mandarin has four
tones, plus a 'neutral' tone. The classification system set up by Yuen Ren Chao in 1930 has been widely used to describe the tones. By this system, the speaker's range is divided into five pitches (numbered one thru five), of which 1 is the lowest, 5 is the highest. The four tones are then represented as follows:

- **first tone**: 55
- **second tone**: 35
- **third tone**: 315
- **fourth tone**: 51

This traditional view has been tested since with acoustic measurements, and confirmed. (C.C. Cheng, p. 40).

Pulleyblank has a different solution to the issue of tone description. Since we are working within the parameters of generative phonology, he claims, we must describe tone by a series of distinctive features. He begins with Nancy Woo's 1969 solution -- that there are two features for tones -- [±High Tone], and [±Low Tone] -- that operate on the segmental level. Contour tones, then, like the Chinese third tone, are made up of sequences of level tones. (Pulleyblank, 1984, p. 37).

Pulleyblank goes on to put forth Moira Yip's generative theory, which extends Goldsmith's (1976) Autosegmental Phonology to account for tones. By this theory, tones are not attached to segments, but form their own tier, from where they are associated with the syllabic unit. Yip goes on to propose her own feature system for tones, which is binary, hierarchical, and confusing, and which Pulleyblank rejects outright. His solution is to combine Yip's autosegmental analysis with Woo's simpler distinctive feature system.

Using this system, tones 1, 2, and 4 can be analysed as [±high tone, ±high tone]; [-high tone/-low tone, ±high tone]; and [±high tone, ±low tone], respectively. For convenience, we can abbreviate the features to HH, MH, and HL. Tone 3, however, gives us a headache. Although it is a contour tone, it is usually realized as the so-called 'half-third tone,' which doesn't have the initial drop, or the final rise. This tone, argue Yip and Woo, is underlyingly simply LL. Yip accounts for the "rise in prepausal position" (Pulleyblank, 1984, p. 57) by a rule that inserts a H tone after the
LL when no other tone follows.

The Mandarin rule that changes a second tone into first tone after first or second tone and before any but neutral is a rule that points to the advantages of having level tone features making up contours rather than having contour features. If we claim that contour tones are made up of contour features, the rule cannot be easily seen as assimilatory, whereas with level features, the high element of the rising or high tone causes the following low element of the rising tone to assimilate.

The TS rule that changes third tone to second before another third tone is an example of dissimilation that also supports the idea of having level tone features. If we describe the third tone as a sequence of LL, the rule can be seen as dissimilating the second L of the first third tone to a H. However, while this rule accounts for a sequence of two third tones, the issue gets more complex when many third tones occur together, as we shall see later.

Hyman and Schuh (1974), make some definitive assumptions about tone universals. Their study is, unfortunately for this paper, based entirely on African tones, but the implications for Chinese are interesting. The conclusions of their article are as follows.

First, a high tone raises the general height of the tones surrounding it -- thus a low tone after a high tone will be realized higher than a low tone following a low tone. Additionally, the stressed tone in a word tends to wrench the tonality from its neighboring segments, leaving them toneless. Contour tones are often formed by tones spreading -- a high tone will last longer than its proscribed time, for example, and spread onto a following low tone, creating a falling tone. And, somewhat opposed to this rule is the tone simplification rule, which deletes these contour tones. By this rule, contour tones simply tend to be levelled.

As stated before, these rules are derived from information taken from African tonal languages. Can they also be applied to Chinese? More importantly, are they really tonal issues at all?

Tone height is by no means an objective entity. One speaker's high tone would well
be acoustically identical to another speakers low tone, for example. Therefore, a high tone is high only by virtue of the distinction it causes among its neighboring tones. Likewise, the concept of tonal downdrift, whereby the overall height of an utterance decreases by the end of the utterance, does not change the relative values of the tones, but merely imposes an intonational pattern upon the sentence which is entirely distinct from the tonal system.

Further, this question of contours spreading and deleting, while it may have been an issue in the development of Chinese tones, is not applicable now, as Chinese tones are well-developed, specified in the lexicon, and fairly rigid in their structure.

Stephen Anderson agrees that downstep does not have to be seen as a problem in distinguishing tones. In a two-tone system such as Hausa, he says, "a high tone may be followed by a tone at the same level, or by a lower tone, and if these are the only possibilities, we can simply label the higher of the two possibilities phonologically as 'high,' regardless of its phonetic value or relation to other tones similarly labelled." (Anderson, p. 141). This neatly sums up the response to the problems posed by downdrift. Although Chinese is more complicated than the two-tone system of Hausa, the same type of argument applies.

The Domain of Tone

Leben, 1978, argues that the domain of tone is a suprasegmental domain. There exists a level above the segmental level, on which tone is mapped. The rules for doing this are straightforward -- first, give the tonal representation the same boundaries as the segmental representation; second, assign tones to vowels in a one-to-one, left to right mapping.

Leben's first proposal is "at the phonological level, a language may have tone patterns expressed for individual words without specifying which parts of the tone pattern are associated with which parts of the word." (Leben, 1978, p. 179)

The above is probably a redundancy proposal for Chinese, since tones are assigned to monosyllabic words in the lexicon before those monosyllables are combined.
There is one aspect of tone in Chinese that is much studied and often debated. When two third tones occur consecutively, the first is read as a rising tone. This is Chinese's famous 'tone sandhi rule.'

Hockett (1947) posits a fifth tone to account for this phenomenon, claiming that this rising tone is overall lower than tone 2. He does, however, include a stipulation that says when tone 5 and 2 are not stressed, they are difficult to distinguish. This makes his argument tenuous, especially in light of his other tone rules. Tones 1 and 3, for example, are usually level, but on occasion not level. "When /1/ is accompanied by loud stress and followed by /3/ (possibly by /1/), there is sometimes a slight fall in pitch at the end..." (Hockett, 1947, p. 257).

Hockett has since been reviewed and revised; C.C. Cheng cites an experiment by Wang and Li (1967), in which they tested native speakers to see if they could distinguish between true second tone and the rising tone that occurs before a third tone. The two tones were indistinguishable (C.C. Cheng, p. 42).

Using the autosegmental approach discussed in Pulleyblank, the tone sandhi rule can be seen as a dissimilation rule -- LL $\Rightarrow$ LH / _LL. "This suggests that the underlying form of tone 2 may actually be LH rather than MH, giving a very symmetrical distribution of the features." (Pulleyblank, 1984, p. 58).

The Tone Sandhi Rule becomes more complex when we introduce a series of three third tone words. C.C. Cheng goes into this question in some detail, and arrives at the conclusion that, in slow speed, the grouping of words syntactically determines when TS will occur. In other words, the famous TS sentence 'Lau Li mai hao jiu' -- 'Old Li buys good wine,' (each of which words is in the third tone underlyingly) would be analysed in the following manner. The NP 'Old Li' would be grouped together, undergo tone sandhi, and emerge as 2-3; similarly, the NP 'good wine' would go through TS as a unit, so that the end sentence would be '2-3-3-2-3.' However, as the speed of the pronunciation increases, he claims, the tone sandhi rule evolves, and its parameters spread as follows.
Slow
(a) 2-3-3-2-3
(b) 2-2-3-2-3
(c) 2-3-2-2-3
Fast
(d) 2-2-2-2-3

Gradually, each of the first four words changes into second tone. C.C. Cheng takes the evolution one step further, claiming that there is a 'fast conversational speech rule' that applies. By this rule, tone 2 becomes tone 1 when following a first or second tone and being followed by any tone except the neutral tone. So, the final, rapidly spoken sentence reads '2-1-1-1-3.' (C.C. Cheng, p. 48). This rule is, even in C.C. Cheng's analysis, a rule of 'conversational casualness,' and is not explored by any other theorists. We shall, therefore, disregard it and end with his fourth listing, at which point the rapidity of speech has caused sandhi to apply to the entire sentence, and give a '2-2-2-2-3' sequence.

C.C. Cheng's explanation for this phenomenon is as follows. 'Syntactic depths' are assigned to the sentence as per Wang's instructions. These instructions yield the following depths for the sentence.

```
NP / 1
A N
lau 1 li 3
S ~VP ~N mai 2 hao 1 jiu
VP NP
```

In slow speech, then, TS applies only across the most shallow sequences; i.e., across 'lau li' and 'hao jiu.' As speed increases, so does the depth of TS application, first applying to levels 1 and 2 (simultaneously), giving the '2-3-2-2-3' sequence, which agrees with sequence (c), then applying to all three levels at once, yielding sequence (d). C.C. Cheng, however, has a problem here. His rule cannot account for sequence (b) which is, according to Chen (below), the most common way to pronounce the sentence. C.C. Cheng adjusts his rule in the following way to account for (b).

Sequence (a), he claims, still has two adjacent third tones. Applying the general form of the Tone Sandhi rule to this yields the appropriate output. So, C.C. Cheng is forced to write the rule "if the output of the first application of the Tone Sandhi Rule includes a
sequence of low tones at the beginning of a phrase, then reapplication of the rule is obligatory" (C.C. Cheng, p. 52). The problem with this, as we shall see below, is when it is applied to large sequences of consecutive third tones.

The question of tone sandhi in sentences with multiple sequences of third tones is one which Matthew Chen (1986a) addresses in great detail. It vexes him that tones should behave in this way that completely defies phonological analysis, and that a system cannot be developed that could accurately predict tone sandhi in slow or rapid speech. Although one of his articles comes to this depressing conclusion:

> On the face of it, facts about Tianjin T5 appear to call into question just about every assumption linguists have long held about the mode of application of phonological rules and their interaction with grammatical structures and with prosodic organization. Either our database is seriously flawed or woefully incomplete, or the analysis presented here is egregiously wrong...

(Chen, 1986b, pp. 15,16), he nevertheless has some very important insights into the mapping of tone sandhi. He begins his 1986a analysis with the famous 'Old Li buys good wine' phrase discussed by C.C. Cheng, above. Chen gives the following derivations of the sequence of third tones, which are basically the same as C.C. Cheng’s shown above.

(from Chen, p. 26)
Again, the problem with C.C. Cheng's analysis is that his solution can't deal with the reading of (b), which is the most natural reading of the sentence. The problem, as Chen sees it, is that "TS is blocked by a relatively weak juncture between 'mai' and 'hao jiu' ... but, surprisingly, arches over the far stronger grammatical division between 'lao li' and 'mai.'" (Chen, 27). In C.C. Cheng's terms, TS applies to depth 3 before applying to depth 2. Chen doesn't like C.C. Cheng's ad hoc solution to this problem -- 'It is awkward, to say the least, to have an optional rule that duplicates the function of an obligatory (but juncture-sensitive) rule of TS." (Chen, 27). Chen concludes that something is amiss in this analysis.

Chen's solution is to apply the rule cyclically.

\[
\begin{array}{ccccccc}
[lao \text{ li}] & [mai & [hao \text{ jiu }]] \\
3 & 3 & 3 & 3 & 3 \\
2 & 2 & \\
\text{n}_a & \\
\_ & 2 \\
\end{array}
\]

By this analysis, TS applies first to the internal brackets, those that C.C. Cheng calls level 1. After the first application, the rule goes back to look at the next brackets. In this case, the second tone 3 has already become tone 2, so TS cannot apply. The final application of the rule changes 'li' to tone 2, yielding the result that C.C. Cheng could not account for.

Things get totally out of hand for Chen, however, when he expands his database to include a series of 12 tone-3 words, as below.
[This sentence translates roughly as 'I would like to ask you to buy two good umbrellas for me']

To start to classify this sentence, Chen formulates some 'foot formation rules,' as follows:

a. IC (Immediate Constituency): link immediate constituents into disyllabic feet.
   b. DM (Duple Meter): Scanning from left to right, string together unpaired syllables into binary feet.
   c. TM (Triple Meter): Join any leftover monosyllable to a neighboring binary foot according to the direction of syntactic branching.

(Chen, 1986a, p. 31)

Then, with TS applying within the domain of the foot, the sentence above can be analyzed as follows:

<table>
<thead>
<tr>
<th>wo xiang qing ni gei wo mai liang ba hao yu-san</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3</td>
</tr>
</tbody>
</table>

(a) IC

F3
F4
F5

(b) DM

F1
F2

(c) TM

F'4
F'5

(d) 2 3 2 3 2 3 2 3 3 3 2 3

(Chen, 1986a, p. 35)

F3, F4, and F5 are all pairs that are dominated by a single node; they fall into the immediate constituents category. Then, unpaired syllables are paired, giving F1 and F2. Finally, 'mai' and 'hao' are joined to their closest neighbors by syntactic branching,
yielding F'4 and F'5. TS first applies to all the binary feet, yielding (d). Next, TS applies to the trisyllabic feet -- in this case, TS cannot apply since the trisyllabic feet no longer contain sequences of two third tones. (d) is an acceptable slowly-read sequence; however, Chen remarks, readings like the following occur at faster paces (with the changed tone identified by bold type), and our phonology must be able to account for them.

A strictly syntactic structure (one like C.C. Cheng's, that eventually changes all sequences of 3-3 into 2-3) could only yield the result in (c). A postlexical analysis, however, can group phrases in any of the above ways. In examples (a) and (b), TS takes place only within each phrase (f1, f2, etc.), and not across phrase boundaries. So, the postlexical analysis can account for all three examples.
Can we fit an analysis of Chinese phonology into any of various phonological mapping methods proposed recently? To further develop this question, we will look at two analyses, that of Diana Archangeli, and that of Moira Yip, who bases her theory largely on Nick Clements' earlier analysis, and see if they can account for the various parts of Chinese phonology.

Archangeli first goes through a brief critique of autosegmental work. Goldsmith and Leben were two of the first autosegmentalists, and started with a relatively two-dimensional analysis, which had only a segment line and a tonal line. McCarthy's (1981) analysis of consonant melodies in Semitic created a more three-dimensional system -- because in Semitic the "consonants, vowels, and the CV pattern each constitute a separate morpheme" (Archangeli, p. 337), all three layers needed to be represented. In addition to these melodies, then, Archangeli adds a fourth layer, that of the syllable.

Archangeli proposed a theory that she called Co-planar representation, and justified this theory with evidence from Yokuts vowel harmony. According to the coplanar representation, there exists a core skeleton, around which are arrayed various planes, in which are several tiers. By this analysis, a core skeleton would have, for example, a vowel plane, a consonant plane, and a syllable plane. In the case of Chinese, though she doesn't discuss it, there would presumably be a plane for tone as well. On each of these planes, then, would be tiers associated with the features of that particular plane. For example, the features for vowels in Yokuts must be separated onto two tiers -- a [round, back] tier, and a [high] tier. The [round, back] tier is docked into the [high] tier, so that roundness and backness can spread onto the [high] tier, but height spreads only onto the core skeleton.

Archangeli uses this model to account for vowel harmony in Yokuts. This model works well to account for vowel harmony; since vowel features are on a separate plane from consonant features, they can assimilate easily, without any interference from the consonants that intervene on the core skeleton. Can we use this model for Chinese?

The first problem I see is that in Chinese, not only is there no vowel harmony; there is
instead harmony between vowels and consonants, or, in Pulleyblank's (1983) analysis, harmony that causes glides to become vowels. These types of harmony are assimilation of vowels to adjacent consonants, not assimilation of vowels over consonants. Right from the start, then, we can simplify the analysis by getting rid of the vowel melody / consonant melody distinction. If we take Pulleyblank's analysis, and say that underlyingly Chinese syllables have a C(C)VC structure, then we can essentially write this into our syllable tier as a permanent fixture.

If it is indeed the case that vowels never harmonize in Chinese, then we don't need to worry about association lines crossing, and can comfortably put vowels and consonants on the same tier.

Then there is the issue of tone, one which few people have tried to account for in their phonological models. In Chinese, tone is specified in the lexicon, is made up of distinct level features, and belongs to the syllabic unit. If we try to put this into Archangeli's model, we encounter some problems. Tone could possibly be represented on a separate plane, which would tie into only the syllabic units. This plane would have tiers labelled [high] and [low]. The problem, then, would be how to represent tones on syllables, and how to write tone sandhi rules. With Archangeli's primarily binary system, it would make sense to represent tones as combinations of [Ø(high)] and [-high]. By this analysis, the four tones would surface as [ØH], [-H, ØH], [-H], and [ØH, -H], where combinations of features are linear, rather than simultaneous articulations. One problem already is that tones 2 and 4 appear to be longer tones than 1 and 3, which is not the case.

The tone sandhi rule is difficult to represent with this selection of features. The rule would have to be stated as follows:

\[-H \Rightarrow [-H, ØH] / _[-H]\]

This rule can at first be seen as a dissimilation rule, with the [ØH] feature being inserted to prevent the sequence [-H], [-H], but if this were the case, a third tone after a fourth tone would also become second tone, which is also not the case.

It seems clear that tone cannot be easily represented on this schematic. A model
which is similar in some ways is that which was first developed by Clements (1985), and most recently modified by Yip (1988). We will follow Yip's analysis, as it seems the most complete. The basic model of this theory looks like this.

This schematic differs from Clements' in several important ways. Clements includes nasality in the manner tier, though he says he doesn't really have a motivation for doing so, nor does he have a major objection to it. Yip solves the problem more neatly, it seems, by including a separate node under the supralaryngeal. Nasality is, strictly speaking, not a mannerism. Further, Yip does not include nearly as many tiers under the manner node as Clements does. It's not clear what she's done with features like [consonantal, sonorant, lateral, strident], however, and perhaps it would be best to include them, as Clements did, in the manner tier.

Additionally, instead of separating place tiers into primary and secondary features, as Clements does, Yip provides labial, coronal, and dorsal nodes. With this system, the vowel features, which are on the dorsal and labial nodes, can assimilate without interference from consonants in the same way they did with Clements' feature system, only this way of classifying them seems much more natural. Additionally, this classification is much clearer and more encompassing than that of Archangeli, because, instead of separating a vowel plane from a consonant plane, in a system which wouldn't allow for adjacent consonants and vowels to undergo assimilation, the features themselves are on separate tiers, and are allowed to spread to consonants or vowels.
With this system, then, vowel harmony can only occur in a language for which the features [high] and [back] are unspecified for consonants. There is no [-feature] in this system, so if the feature is unspecified, the node disappears; therefore, if [high] and [back] are unspecified for consonants, the consonants do not show up on the dorsal level, and vowels that want to harmonize are adjacent on that level.

Again, features for tone are not placed on the schematic, although Yip does comment briefly on the nature of tone. Tonal features, she explains, are usually also unspecified for consonants, so non-adjacent syllabic slots would be adjacent on the tonal tier. Tone is not a place feature, so it would have to form a separate node under the supralaryngeal.

Later, Yip claims that the Tone node is a superordinate node, like that of Root. I'm not sure exactly what she means here....

But, since there are no [-feature] in this system, we can return to our classification of tones by the [high] and [low] features. Thus, the tones are again [HH], [LH], [LL], and [HL]. The dissimilation rule of tone sandhi, then, looks like this, in conventional notation.

\[ [LL] \Rightarrow [LH] / \_ \_ [LL] \]

This rule at least makes sense, in that it prevents a sequence of four L features. Of course, this kind of phonological rationale cannot account for the idiosyncracies of the TS rule when it applies to long sequences of third tones, but it does solve the basic TS problem. Additionally, it seems that the kind of notation that Yip et al. have used does not extend well to dissimilation rules. So, in order to write the rule in their notation, we would have to come up with some extra maxims.

In order to account for the change, each syllabic unit must have possible ties to one H
and two L features. H and L features, however, would each have their own tier, so that the four consecutive L's could see each other. The rule, then would call for dissociation of the second L feature, and a rule that would require each syllabic unit to be tied to two tone features would associate the syllabic unit to a H feature by default.

This theory seems implausible at best -- but there seems to be no other way to account for dissimilation processes in this model.

In a recent conversation with Nick Clements, a more plausible solution was suggested to me. This theory would claim that tone three has an underlying representation of LLH. In order for this to work, we must include the maxim stated above, that each tone-bearing-unit must be attached to only two tone features. The LLH sequence, then, is realized in ordinary speech as LL, and only surfaces fully when the word occurs in isolation, or at the end of a sentence. Then, for TS to take place, the second L of the first third tone delinks, and the H, which was floating until then, is attached. The impetus for TS here, then, is the sequence of four L tone features. The phrase 'lau li' would look like this.

\[
\begin{align*}
\text{lau} & \quad \text{LL} \quad \text{L} \\
\text{li} & \quad \text{L} \\
\end{align*}
\]

This theory definitely seems the most plausible -- it accounts for the behavior of the third tone in isolation, as well as providing motivation for TS. It also solves the problem of writing this particular dissimilation rule within the constraints of the model. This leaves open the possibility that in all dissimilation processes, the opposite feature is actually present in the UR, and only surfaces under certain conditions.
References


