Tonal features, tonal inventories and phonetic targets[*]

MOIRA YIP

Abstract

Most work on Asian languages uses two binary tone features. There are however cases in which the Register System cannot accurately characterize the phonetics of the recorded tones, particularly mid and contour tones. This paper identifies these problems, and sketches solutions. It proposes a new algorithm for converting featural specifications into phonetic levels. Next, tonal inventories are analyzed within an OT instantiation of Dispersion Theory. Finally it suggests that contours have only an initial tonal specification, and the phonetic contour is the result of drift away from that initial target. Level tones have two targets, firmly anchoring both ends and keeping the pitch stable.

1 Introduction

In the last 20 years there has been a large body of work on Asian tone languages that uses a four-level feature system first proposed in Yip (1980), and subsequently modified by a variety of authors including Pulleyblank (1986), Bao (1990), Duanmu (1990) Hyman (1993) Chen (2000) and others. I will refer to this as the Register System. However, there are cases in which the Register System cannot accurately characterize the phonetics of the recorded tones.

There are three problems related to mid tones. First, the model is a poor fit for three-toned systems in which the mid tone is phonetically equidistant from both H and L, and demonstrably not the unspecified default tone. Second, some languages appear to treat featurally distinct but phonetically identical mid tones as identical even in some aspects

[*] Many thanks to Jaye Padgett, Edward Flemming, Mark Liberman, Jonah Liu, Robert Kirchner, Dan Silverman, Jie Zhang, and audiences at ZAS in Berlin, and at UCL. Special thanks to two anonymous reviewers. Conversations with all of you greatly shaped this paper. It was made possible by a generous grant from the Chiang Ching Kuo Foundation, and also by the hospitality and support of University College London, particularly Neil Smith and John Harris. All errors, inaccuracies, and omissions are of course my own responsibility.

of the phonology. Third, three-tone systems are always H,M,L, with the gap being a missing mid tone, not a missing H or L.

A different problem concerns the phonetics of contour tones. Many, perhaps most, do not have the shapes one might expect in this model, which analyzes them as sequences of two level tones, each of which should therefore be a phonetic target, so that a rise looks like this: [_/¯].

Three main proposals will be made in this paper. First, a specific algorithm for converting the Register System's featural specifications into phonetic levels will be proposed that will solve the first mid-tone problem. Second, the combinations of tones found in tonal inventories will be analyzed within an OT instantiation of Dispersion Theory (Lindblom (1990), Flemming (1995)), which will explain why three-tone systems are missing a mid, not a high or a low, and why contour tones are so widely used. Third, it will be suggested that contours have only an initial tonal specification, and that in the phonetics the contour is the result of drift away from that initial target, whereas level tones have two targets, firmly anchoring both ends and keeping the pitch stable. The data are drawn entirely from Chinese, and it remains to be seen to what extent data from other language families bear out these still speculative proposals.

§2 sets out the problems in more detail, §3 looks at the first two mid-tone problems, §4 looks at inventories, and §5 looks at contour tones.

2. Two problems for the Register System tonal feature set

2.1 The mid-tone problems*:*

Yip (1980), Pulleyblank (1986), Bao(1990), Duanmu (1990) and others have argued for a set of two binary features, one of which, [+/- Upper], divides the pitch range into two registers, and the second of which, [+/-high], fine-tunes this. Below I show such a system as it might map onto systems of three or four tones. The lowest tone in the system nearly always shows a slight fall, as in [21], but I follow most previous work in analyzing such tones as phonologically level (Yip 1980, Chen 2000). In this paper I use the system of Chao (1930) in which 5 means highest tone, and 1 means lowest tone.

(1)

		4-tones	*3-tones*	
			Option 1	*Option 2*
[+ Upper] (+U)	[+high] (h)	55	55	55
	[-high] (l)	44	33	
[-Upper] (-U)	[+high] (h)	33		33
	[-high] (l)	22/21/11	22/21/11	22/21/11

Importantly, in a three-tone system reported as {55, 33, 11} the mid tone could in principle be either [+U, l] or [-U,h], as shown in the two right-hand columns above, but it must be phonologically one or the other, implying that it is either closer to the highest tone, or the lowest tone, but not equidistant. The only apparent alternative is to consider it unspecified for all tonal features, with its value filled in at the phonetic stage. In that case it might be truly intermediate between high and low tones in isolation, but it should be contextually variable. We thus encounter our first problem with this feature system. Under specification aside, mid tones are forcibly assigned to one or other register, even if the phonetics do not justify this. If we instead insist on taking the phonology to phonetics interpretation seriously, any mid tone that is phonetically equidistant from high and low tones can only be explained as phonologically unspecified. But in that case it has no phonological target, and its exact pitch should be contextually variable. However, Taiwanese, §3.1, shows that a language can have a invariant mid-tone target that is exactly equidistant from the high and the low tones, something this feature system cannot directly represent.

The Taiwanese data suggest that we must loosen the link between the features and the phonetic pitch. Given this looser link, we can then also imagine a language with two mid tones that are phonologically distinct, but phonetically identical. Such languages do exist, but unfortunately some aspects of the phonology seem at first glance to pay attention less to the demonstrable phonological distinction between these tones than to their phonetic similarity. In §3.2 I will illustrate this problem with data from Wenzhou and Cantonese.

Finally, we must explain the odd fact that tonal inventories with fewer than four level tones are usually {H, M, L}, with only one mid-level tone, rather than , say, {M+, M, L}, (where M+ stands for a higher-mid tone), with the H tone missing. Given two binary features, we might expect to find languages with any one of the four possible levels missing, and yet that does not seem to be typical: the missing tone is usually one of the two possible mids. In §4 I outline a Dispersion Theory approach to this.

2.2 Contour tones: a one-target proposal

Linguists have argued forcefully that contour tones are sequences of level tones in most if not all languages (see for example Andersen 1978). The feature system given above, together with the assumption that register, [Upper], is held constant across the syllable, allows for two rises and two falls, mapped onto a four-contour system as shown in the first two columns below. If there is a three-contour system with only one rise, such as [24], the mapping has two options, shown in the last two columns below: the rise could be [+U, lh] or [-U, lh].

(2)

		4-contours		*3-contours*	
				Option 1	*Option 2*
[+Upper]	hl	53	42	53	53
	lh	35	24	24	
[-Upper]	hl	31	31	42	42
	lh	24	13		24

When we consider the phonetics of contours, we might expect each tone to be realized as a target, with interpolation in between (Pierrehumbert and Beckman 1988) giving a shape like this for a LH rise: [_/¯], with a plateau at each end of the contour. In fact, however, the overwhelming majority of contours, in Asian languages at least, have an initial plateau only, so that a fall looks like this: [¯\ ] and a rise looks like this: [_/]. This suggests an alternative view. Contours are tones with a target specified at the start of the syllable, and then they drift away from that target (while staying within register). Level tones are tones with targets specified at both ends of the syllable, and the tone thus remains stable throughout. This proposal will gain support from the phonology of Cantonese in §3.2, and from the phonetics of Mandarin tones, §5.1.

3. The first two mid-tone problems: a closer look

As I have said, there are three problems concerning the representation of mid tones in the standard register feature system. This section deals with the first two in turn, and the third is discussed in §4.

3.1 Phonetic evidence that phonologically specified mid tones in a three-tone system may be equidistant between high and low tones

In this section I show that the mid tone of a three-level system may be a phonetic target just like the high and the low tones, and yet be equidistant from both. The two available featural specifications are [+U, l] and [-U, h], which should translate into a tone closer to either [+U, h] or [-U, l]. To resolve this puzzle, I propose a new algorithm for the realization of tonal features.

Taiwanese (Peng 1997) has three level tones /55, 33,21/. The notation obscures the fact that all three tones fall through about 25Hz phrase medially, and I take them all to be phonologically level. Like all Min dialects these tones change contextually. If one of these tones appears domain-finally, it surfaces unchanged. In all other positions, it changes into a different tone in the system, as shown in the following diagram:

(3) 51 → 55 → 33 ← 24
↖ ↙ Min "tone circle"
21

The $F_{\text{Ø}}$ plot in Figure 1 is taken from Peng. If we take the stable plateaux portions of the 55 and 33 level tones, and extrapolate the [21] tone, whose later parts are obscured by the heavy glottalization typical of very low tones, and then draw a vertical line through the pitch tracks at this point, about 180ms into the syllable, we see that the mid tone in final position is rather precisely poised between the high and low tones:

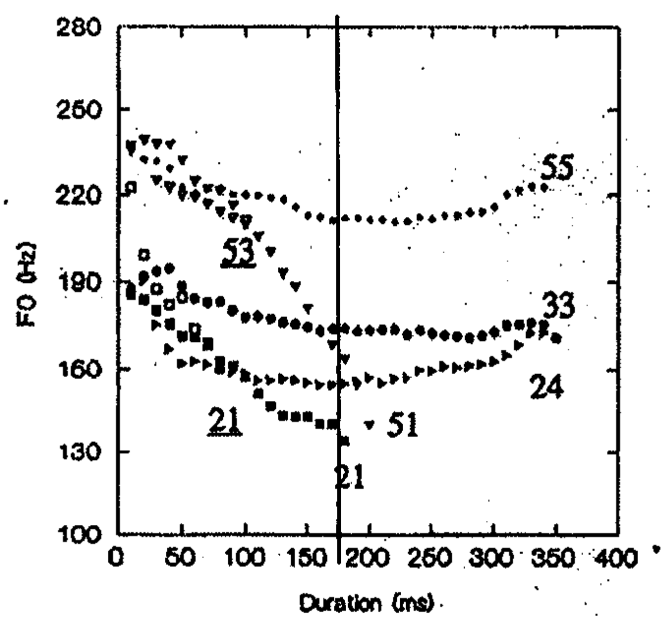


Figure 1: A sample of Taiwanese tones produced by one native speaker.

Measurements of the apparent targets or stable portions of medial high 55, mid 33 and low 21 tones, for four different speakers M1, M2, F1, F2, are extracted into the following table. In addition to the $F_Ø$ values for each tone, I give the differences between them. The final row of the table is the differences of the differences. If the mid-tone is equidistant, each cell in this row will be zero. The reader can see that indeed it is zero for two speakers, and only 5Hz for the other two, showing that the mid tone is truly mid phonetically. Real confirmation of this claim would need a full statistical analysis of Peng's data, since he does not directly address this issue. My conclusions are thus tentative only.

(4) $F_Ø$ values in Hz. for stable parts of three phrase-medial level tones for two male and two female speakers

	M1	F1	M2	F2
55	190	270	190	270
33	160	230	175	230
21	135	190	160	195
(55-33)	30	40	15	40
(33-21)	25	40	15	35
(55-33)-(33-21)	5	0	0	5

In the Register system laid out in (1), I have suggested that this truly mid mid-tone can only be achieved by assuming that it is unspecified for all tonal features, thus avoiding a commitment to [+U] or [-U]. In that case it should be more subject to variation than the other level tones, but further data from Peng shows that this is not the case. Figure 2 gives the pitch track for the medial 33 tone before each of three following tones, [55, 21, and 33]. All three pitch tracks overlay each other and are identical, showing that the following tone has no significant effect on the preceding [33]. This suggests that it must come supplied with its own phonetic target, and therefore with distinctive features.

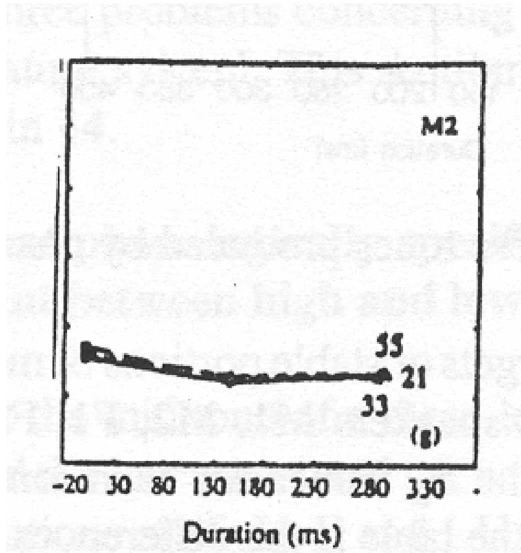


Figure 2: 33 tones before 55, 21 and 33

What then do we conclude from Peng's Taiwanese data? If we assume that an invariant tone implies a phonologically specified target, then either our feature system must allow for stating a mid tone that is *featurally* equidistant from both H and L, or the phonetics must allow a featurally [+U] or [-U] tone to be phonetically realized at the exact boundary between [+U] and [-U] registers.

There are certainly alternative feature systems that have the first property. For example, Hyman (1993) defines the features H, L as follows:

(5) H: At or above a neutral reference height
L: At or below a neutral reference height.

M tones, being at the neutral reference height, are thus simultaneously [H, L]. Alternatively, if one uses binary features analogous to vowel height features, mid tones might be [-high, -low]. The problem is that a large body of work on Asian languages (e.g. Bao 1990, Duanmu 1990, Chen 1996) has shown the utility of the register feature system with which I began this paper, and in particular that the mid tones which are part of contours are clearly associated with the same register as the rest of the contour, so that the mid of a MH is [+U, l] and the mid of a ML is [-U, h].

Sticking to our original feature system, then, I will instead explore an explanation located in the phonetic interpretation algorithm. I propose the following three principles of phonetic interpretation.

(6) (i) The pitch space is exhaustively and equally partitioned into registers.
(ii) Within a register, a [h] is realized at the upper limit of that register, and a [l] is realized at the lower limit of that register, *unless*..
(iii) contrast preservation requires that two tones be kept distinct.

These principles will give us the following three and four tone systems respectively; the dashed line marks the register boundary.

(7)		*Two three-tone systems*			*Four-tone system*
		Option 1		*Option 2*	
		+Upper, high		+Upper, high	+Upper, high
				+Upper, low	
	--	+Upper, low	-----	- Upper, high	-----------------------------
				-Upper, high	
		- Upper, low		- Upper, low	-Upper, low

Principle (i) ensures that the lower boundary of the upper register and the upper boundary of the lower register coincide. In a three-tone system, principle (ii) ensures that either of specifications [+Upper, low] or [-Upper, high] will be realized at the register boundary, and therefore equidistant from H and L. In a four-tone-system principle (iii) will keep [+Upper, low] and [-Upper, high] away from each other, and therefore away from the register divide.

We can now explain the first of the mid-tone puzzles. Mid-tones in a three tone-system are exactly mid because of clause (ii) of the principles of phonetic interpretation. The Taiwanese data are thus accounted for.

3.2 [+Upper, l] and [-Upper, h] mid tones treated as phonologically identical:

In this section we will see two cases in which processes - in one case an OCP-triggered movement, and in the other poetic rhyme - appear to treat mid tones as identical even when they must be featurally distinct, with one being [+Upper, l] and the other [-Upper, h].

3.2.1. Wenzhou. Like most Wu dialects, Wenzhou has a register system with lower variants of all tones after voiced onsets. So we get [taHM] '*dawn*' but [da ML] '*talk*'. Numerically, Chen gives these falls as 42 and 31 respectively. Featurally, these must be [+Upper, hl] and [-Upper, hl] respectively, assuming, with Chen, that contours are sequences of levels. It follows that the M of HM is [+Upper, l] and the M of ML is [-

Upper, h]. The following discussion takes as its starting point a very complete analysis in Chen (2000:475ff).

Unusually for a Wu dialect, Wenzhou has a right-dominant tone-sandhi system. In compounds, all but the tones of the last two syllables are deleted. The final tone remains unchanged. The penultimate tone remains in situ but changes in synchronically messy but historically intelligible ways that affect the first of any two adjacent tones, and depend on the particular tones involved. For example, /wenL-tiML/ '*question*'→ [wenHM-tiML] ; /ziMH.danL/ '*bullet*' → [ziHM.danL]; /tianML.jiM/ '*frog*' → [tianL.jiM]. The reader is referred to Chen for full details.

If the two final syllables are HM.ML after disyllabic tone sandhi has applied, and if the compound is tri-syllabic or longer , instead of staying put the penultimate HM tone moves leftwards away from the final ML to the first syllable across any number of toneless syllables. The movement, which Chen calls HM Shift, is shown schematically below, followed by an example and a derivation. The derivational format, taken from Chen, is purely expository, since in general I will be working in non-derivational Optimality Theory.

(8) [σ......σ σ]$_{word}$ → [σ......σ σ]$_{word}$

```
            |    |              \___     |
            T    T                  T    T
            /\   /\                 /\   /\
            HM.ML                   HM  ML
```

(9)

```
      wireless    telephone tube
    [[wu-xian]-[dian-hua]]-tong       'radio receiver'
      ML  HM     L      L      ML     base tone
    Ø     Ø      Ø      HM     ML     Disyllabic tone sandhi,
                                      and deletion of  all but last two tones
    HM    Ø      Ø      Ø      ML     HM Shift, separating two M tones
```

The relevance of this data to the mid-tone issue is that here we have a rule, HM Shift, that appears to be triggered by the OCP (although Chen does not make this move), and which forces the separation of two M's, but where those M's are featurally distinct. As noted earlier, the M of HM, [+Upper, hl], is [+Upper, l], whereas the M of ML, [-Upper, hl], is [-Upper, h]. No phonological rule should treat them as identical.

There is a possible solution available: perhaps the tonal movement is triggered by the OCP's recognition of a sequence of identical [hl] contours. Both HM and ML are featurally [hl] falls. There is some supporting evidence for this view. First, /HM.MH/

inputs survive unchanged, so an analysis with the OCP acting on M.M sequences is unlikely. Secondly, there are known to be other cases of the OCP operating on full contours. In Tianjin, the OCP causes leveling on the first syllable of two identical contours, so that LH.LH → H.LH and HL.HL → L.HL (Yip 1989). In Pingyao, the OCP causes contour reversal in similar circumstances, so that HM.HM → MH.HM and MH.MH → HM . MH (Bao 1999). Phenomena like these are used by Yip and Bao as evidence for the grouping of tonal sequences into a constituent under a tonal node, with the OCP acting on this node. Given this alternative, I conclude that Wenzhou is not a clear-cut problem for the representation of mid-tone. A more serious problem is posed by Cantonese, to which I now turn.

3.2.2. Tonal identity in Cantonese song:

Cantonese has six tones. Data come from Chan (1987).

(10)	55	[+Upper, h]	22	[-Upper, h]
	33	[+Upper, l]	11	[-Upper, l]
	35	[+Upper, lh]	13	[-Upper, lh]

[55] has a [53] variant for some speakers. [11] is often transcribed as [21], as in so many dialects. Hashimoto (1972) gives both level and falling as free variants. I shall assume that the initial fall is phonetic. [13] is sometimes given as [24]. Phonetically, the ending pitches of 33 and 13 are very close. Featurally, the [33] tone is the second highest of four level tones, and therefore must have the features [+U, l]. The [13] tone is the lower of two rises, and must have the features [-U, lh].

In modern Cantonese songs with several stanzas, each line has requirements on the tones that can occur on each syllable. The first line of one song must have the following tones in each stanza: 13 and 33 are interchangeable, as are 55 and 35.

(11)	$\sigma_1\sigma_2\sigma_3$	σ_4	σ_5	σ_6	σ_7	σ_8		
	22	11	22	13/33	55/35	55/35	13/33	55/35

The pairs 13/33 and 35/55 are clearly paired because they share an end-point. This is easily stateable in phonetic terms, but, for the 13/33 pair, not in featural terms since [13] is [-U, lh] and ends on [h], and [33] is [+U, l]. Note that the remaining tones 22 and 21 cannot pair with each other, or with any other tone. [1]

The solutions to this problem are not obvious. One could perhaps say that songs pay attention to the phonetics of the tones, and tell us nothing about the phonology, but this goes against a body of research showing that folk poetry and song in many languages pay attention to a more abstract level of representation than pure phonetics (Mohanan 1986, Fabb 1997). Alternatively one could appeal to under specification of some sort. Suppose that the 33 tone is the absence of features, and that the 13 tone is /[-U, l], Ø/, a tone that has a low initial target on the first mora, and then no later target on the second mora. We then say that the slot in the tonal scheme for poetry cannot be filled by anything with a final l or h target, leaving only 33 and 13 as possibilities. This proposal will receive some support from Mandarin in section 5.1, where the representation of the contour tones is given in full.

Having shown that the phonology does not always treat mid tones consistently, I now turn to the last of the three mid-tone problems: why inventories with three tones always lack one of the mid tones, not high or low.

4. A Dispersion Theory approach to inventories: Why mid is mid, and why contours are needed at all:

I now turn to the remaining unresolved mid-tone problem. Why do three tone systems have high, low and mid rather than, say, high, raised-mid and mid? In other words, why is the inventory gap always a mid tone rather than a low or a high tone? I shall suggest that the answers lie in the pressures placed on inventories by the demands of the articulatory and perceptual systems, particularly the latter.

4.1 Chaoyang tonal inventories

I begin by roughing out a theory of tonal inventories, using Chaoyang, a Min dialect, as my example. (Zhang 1979, 1981, 1982). Like Taiwanese, Chaoyang shows a chain shift effect whereby every tone changes to a different tone from the basic inventory when in non-final, i.e. non-head, position. The changes are not easily stated in any natural phonological way, since a tone that is pre-pausal for one syllable may be the non-final

---

[1] A reviewer suggests that the [11/21] tone might be phonologically falling, [-U, hl]. In that case [33] can be [-U,h] and [22] can be [-U,l]. The 33=13 rhyme now falls out as rhyming the two tones ending in [-U,h]. However, a new problem arises - why don't 22 and 11/21 rhyme, since both end on [-U,l]?

variant for a different syllable. This is because, as in Taiwanese, the changes create a chain. Look at the following data, showing only sonorant-final syllables. The tone in the first column is the citation, pre-pausal form, and the one in the second column is the non-final variant. So a pre-pausal /53/ becomes a non-final 31, but a pre-pausal /31/ becomes a non-final 55, and so on. The final column shows what happens when a syllable is completely unstressed.

(12) *Chaoyang tone inventories in different contexts:*

<table>
<tr><th colspan="2">Final/Citation
= Head</th><th colspan="2">Non-final Sandhi
=Non-head</th><th colspan="2">Post-tonic,
unstressed</th></tr>
<tr><td>55</td><td>H</td><td>11</td><td>L</td><td rowspan="5">11</td><td rowspan="5">L</td></tr>
<tr><td>31</td><td>ML</td><td>55</td><td>H</td></tr>
<tr><td>33</td><td>M</td><td rowspan="3">33</td><td rowspan="3">M</td></tr>
<tr><td>313</td><td>LM</td></tr>
<tr><td>11</td><td>L</td></tr>
<tr><td>53</td><td>HM</td><td>31</td><td>ML</td><td>31</td><td>ML</td></tr>
</table>

At first glance these changes are mysterious phonologically. It has been suggested for the closely-related Min dialect Taiwanese that they may not be productive: psycholinguistic experiments by Hsieh (1976) and Tsay and Myers (1996) suggest that if speakers are given a nonce form in citation tone, and then asked to use it in a sandhi context, they answer more or less at chance. Tsay and Myers thus suggest that all lexical items come in pairs, and that a lexical insertion rule controls the choice. The problem here is that any pairing could be learnt, but in fact all words, including loans, conform to the standard pairings of allotones. Sometimes the changes are more easily stated historically (Chen 2000). I shall explore a partially phonological account, based on the observation that usually (but not always) there is some degree of neutralization in non-final position (leading most researchers to take the citation tones as basic), and, more importantly, that the inventory of tones in each position is progressively less marked as the stress reduces.

To begin, note that markedness cannot explain the individual tonal changes. For example, in Chaoyang, suppose 55 becomes 11 in non-head (non-final) position because 55 is marked and 11 is unmarked. Then we must ask why 31 becomes the ‘marked’ 55 in the same context and why the ‘unmarked’ 11 becomes the presumably more marked

33. However, as I have said, the non-head inventory as a whole is usually less marked: it has fewer tones, and the tones that are lost are typically marked, especially rising tones. I propose then that the individual pairings are indeed lexical, but they must be picked from the inventories appropriate for the position in question. What remains for the phonologist is to explain the particular inventories found in each position.

The following markedness facts will play a role; all of these are clearly underpinned by a tendency to minimize articulatory effort, and can be thought of as constraints:

(13) *Minimize Articulatory Effort* (MINARTICEFF)

a. Contour tones are more marked than level tones: *CONTOUR
b. Rising tones are more marked than falling tones: *RISE>>*FALL
c. High tones are more marked than low tones: *H >> *L

The justification for these constraints rests on well-known observations about tonal markedness, illustrated here with examples from Chinese. Hashimoto (1980) surveys 83 dialects of Chinese. She looks at the changes that take place in tone sandhi, typically a site of neutralization, and notes that leveling is the most common change. Similarly in the acquisition of Mandarin, there is a tendency to level off falling or rising tones (Clumeck 1980, Li and Thompson 1978). In inventories, however, level tones are less common than falling ones (Cheng 1973). Hashimoto attributes this to pre-pausal effects, since citation tones on the monosyllabic morphemes of Chinese are of course pre-pausal in isolation.

Changes to falling are very common. Changes to rising are also found, but they are much rarer, and are often the result of dissimilation. In acquisition, rising tone tends to be acquired last. Rising tones take longer to produce than falling tones (Sundberg (1979), Zhang 1999), and they are less common in inventories overall.(Cheng: 1973).

Lastly, in many languages, especially in Africa, phonologically toneless syllables surface with unmarked low tone, and lowering of pitch is common in Chinese in sandhi contexts. When raising of tone is found, it is usually as part of a flip-flop or chain-shift pattern as in Min or Hakka.

These articulatorily-based constraints are balanced by a set of perceptually based ones. Lindblom (1990) observes that inventories tend to spread out to occupy the available phonetic space, with clear functional advantages. Applying his Dispersion Theory to tone, inventories with few tones have those tones as widely spaced as possible at the edges of the pitch space and inventories with more tones have them closer together, but still regularly spaced in the tonal space. Flemming (1995) develops this idea in OT for segmental inventories, and the following account is heavily influenced by his work. See also Jun 1995, Silverman 1995, Kirchner 1998, Steriade 1995, Hayes 1996, Padgett 1997, Zhang 1999.

In what follows you will see tableaux that compare tonal inventories, not different output candidates for a given input. Suppose that Lindblom's Dispersion Theory is formulated as space constraints that require spacing the tones some given percentage of the pitch range apart. Note that I assume that Space constraints, being perceptually based, apply only to tones of the same shape, most particularly level tones, and therefore distinguished by pitch alone.

(14) SPACE-100%: Level tones must be 100% of the pitch range apart.
Can only be satisfied by systems with 2 or fewer evenly spaced tones:
{H, L}, or any single tone

(15) SPACE-50%: Level tones must be 50% of the pitch range apart.
Can only be satisfied by systems with 3 or fewer evenly spaced tones:
{H,M,L}, {H,M}, {H,L}, {M,L}, or any single tone

SPACE-50% is always ranked above SPACE-100%,and in general SPACE-n% is always ranked above SPACE-(n+1)%. See Flemming for discussion of this issue as it pertains to vowel formants.

It is these constraints that enforce equal spacing of level tones, and thus pick out a true mid tone equidistant from both H and L. Only if this constraint is outranked by some other constraint, such as a prohibition on L tones in head syllables (see De Lacy 1998) will we get a skewed inventory such as /H, M/. Different languages have different numbers of tones, so this must be stipulated. In Chaoyang, the number changes with prosodic context, from 6T to 4T to 2T, instantiated here as constraints encouraging maximizing the number of contrasts:

(16) 6T/4T/2T: Distinguish 6/4/2 tonal distinctions

Universally, 2T>>4T>>6T, or, more generally, nT>>(n+1)T. Given the spacing requirements, sufficient distinctions for a rich tonal system are not always possible with level tones only. *RISE exacts a price for adding rises, so if possible, falls are added instead. *H exacts a price for adding H, so any contours will if possible be ML or LM.

Let's start with the smallest inventory of Chaoyang, {L, ML} in post-tonic position. This inventory is the perfect way to achieve a two tone inventory, while having no H, no rises, and keeping the level tones at least 100% of the pitch range apart. The high ranking of SPACE-100% is undoubtedly related to the brevity of the unstressed syllable, and therefore the difficulty of perceiving FØ distinctions. The tableau tells the tale. Candidate (a) has insufficient tones, candidate (b) has them too closely spaced, candidate (c) violates *RISE, and candidates (d) and (e) violate *H. The winner is thus (f), with the low falling ML tone.

(17) Unstressed inventory

	2T	SPACE-100%	MINARTICEFF	
			*RISE	*H
a. L	*!			
b. L, M		*!		
c. L, LM			*!	
d. L,HL				*!
e. L, H				*!
☞ f. L,ML				

In sandhi positions there are 4 tones {H,M,L,ML}, and the SPACE-100% requirement has been demoted off the right-edge of the tableau, and is not shown here. Instead, the less stringent SPACE-50% plays a role in limiting the number of level tones to three, and forcing them to be equally spaced H,M,L. The four-tone inventory is achieved by adding the least marked contour tone, a fall from ML that includes no H:

(18) Sandhi Tone Inventory

	SPACE-50%	4T	MINARTICEFF	
			*RISE	*H
☞ a. H, M, L, ML				*
b. H,M,L,HM				**!
c. H,M,L,LM			*!	*
d. M,L,ML,LM			*!	
e. H,M,L		*!		*
f. H, M+, M, L	*!			*

Finally, the six tone citation inventory {H,M,L,ML,HM,LM} still observes SPACE-50%, again limiting the number of level tones to three, so the six tones are achieved by adding three contours. To do this, *RISE and *H must unavoidably be violated, but the minimum violations are achieved with two falls, ML and HM, and one rise, LM:

(19) Stressed syllable inventory

	SPACE-50%	6T	MINARTICEFF	
			*RISE	*H
a. H+,H,M+,M,L+,L	*!			**
b. H,M,L		*!		*
c. H,M,L,ML,MH,LM			**!	**
d. H,M,L,ML,HM,MH			*	***!
☞ e. H,M,L,ML,HM,LM			*	**

Note that the MINARTICEFF markedness ranking for Chaoyang, *RISE >> *H >> *FALL, *L, is invariant. All that changes in context is the relative ranking of MINARTICEFF and the cross-linguistically fixed SPACE hierarchy relative to the cross-linguistically fixed “nT” hierarchy.

The number of contrasts is conditioned by prominence. Although the precise number of tone contrasts in any given language (like the number of contrasting vowels) is language-specific, I would claim that cross-linguistically as prominence decreases, the number of contrasts stays stable or shrinks, but never increases. The reasons for this are fairly clear. Less prominent syllables are often shorter, giving less time to articulate distinctions, and thus encouraging articulatorily “easy” tones. It also means there is less time to perceive those distinctions, and thus it is preferable to use fewer, more widely-spaced tones in order to make the remaining contrasts detectable. In addition less prominent syllables are often quieter, which also makes contrasts harder to perceive, and places a premium on highly distinct contrasts. (Zhang 1999). In OT terms, then, MINARTICEFFORT (articulatory markedness constraints), and SPACE-100% (maximizing contrasts) both move up the hierarchy relative to the “nT” constraints. It is their exact position in the fixed nT hierarchy that determines how many tones the language has in that particular prosodic position. If we compare the three hierarchies motivated earlier, we see the following picture emerge. I have bolded the constraints that get promoted as prominence decreases:

(20) *Max prominence: Citation*
SPACE 50%>>2T>>3T>>4T>>5T>>6T>>**MINARTICEFF>>SPACE100%**

(21) *Medium prominence: Sandhi*
SPACE50%>>2T>>3T>>4T>>**MINARTICEFF>>SPACE100%**>>5T>>6T

(22) *Minimal prominence: Neutral/post-tonic*
SPACE50%>>2T>>**MINARTICEFF**>>**SPACE100%**>> 3T>>4T>>5T>>6T

Once the inventory is defined for a given context, it constrains the particular output by requiring that the tone be part of the inventory set. It does not, however, explain why a given lexical item chooses the tone that it does, and here as I said earlier, I tend to the view that the pairings are indeed lexical, a position I have resisted for years but at present see no alternative to. The interested reader might look at Kirchner (1996) on chain shifts as a possible starting point for a more interesting analysis.

This approach also allows an answer to a different problem raised in section 1. Many systems have contours that use levels the system doesn’t use This is hard to explain in purely featural terms, where one might expect no contours at all until all possible levels are used in solo roles. For example, Mandarin has no level M, but it uses it in contours. Conversely, the level L is not used as a component of contours, but it is used as a solitary level tone.

(23) Mandarin:	55	H	[+U, h]
	21	L	[-U,l]
	53	HM	[+U, hl]
	35	MH	[+U, lh]

We can now view this as the result of perceptual pressures. If space 100% is high-ranked, for example, only two level tones H, L are allowed. So extra contrasts must come from contour tones. They offer a way of building a large tonal inventory without spacing levels too closely.

4.2 Heads prefer H: Mandarin

The particular choice of contours sometimes has a different cause. In general as we have seen ML will be preferred to HL (or MH) as a contour, because of the high-ranked *H. However, *H may in turn be dominated by a requirement for heads to have H tone, HEAD=H, (Zoll 1997, De Lacy 1998). Since in Mandarin all fully-toned syllables are bimoraic heads (Duanmu 2000), most tones include a H, and we get the inventory in (23) above. The full grammar needed to achieve this is 4T, SPACE 100% >> HEAD=H >> *H >> *RISE, *FALL >> *L. In non-heads, HEAD=H will play no role, so *H will step to the fore, and we will get a reversion to unmarked L tone.

5. Phonetic evidence that contours may not be composed of two fixed level targets: 5.1 Mandarin

If contours are sequences of levels, the most straightforward expectation is that each level tone should be a phonetic target. A LH rise should then start from a low plateau, and rise towards a high plateau at the end, looking roughly like this: [_/¯]. In this section I will argue that the Chinese data do not show contours of this shape, but instead appear to have only an initial target from which the pitch then moves away.

Xu and Wang (1999) conducted a phonetic study of Mandarin tones in combination. They argue that the phonetic targets are level H and level L, plus Rising and Falling, and they explicitly argue against analyzing these contour targets as sequences of LH and HL targets. Their evidence comes from the shape of the envelope in co-articulation effects. They set the scene with the following figure. In Figure 3(a), we see a schematic picture of two possible rising tones following a low tone. The upper, solid tonal curve is what one would expect to see if the rise begins with a L target, and then approaches a level H target (shown by the solid horizontal line at the top). The lower, dashed line is what one would expect to see if the rise is a Rising target, shown by the diagonal dotted line in the lower right, and the pitch is asymptotic to this target incline. On the right, in Figure 3 (b), we see the transitions from a preceding H tone to a rise. Again the solid curve shows what we would expect if the rise has first a L target, then a H target: a deep dipping contour. The dashed line shows what we would expect if the rising tone target is the diagonal Rising line: an $F_{\emptyset}$ trace asymptotic to this line, never dipping very low, and with no plateaux.

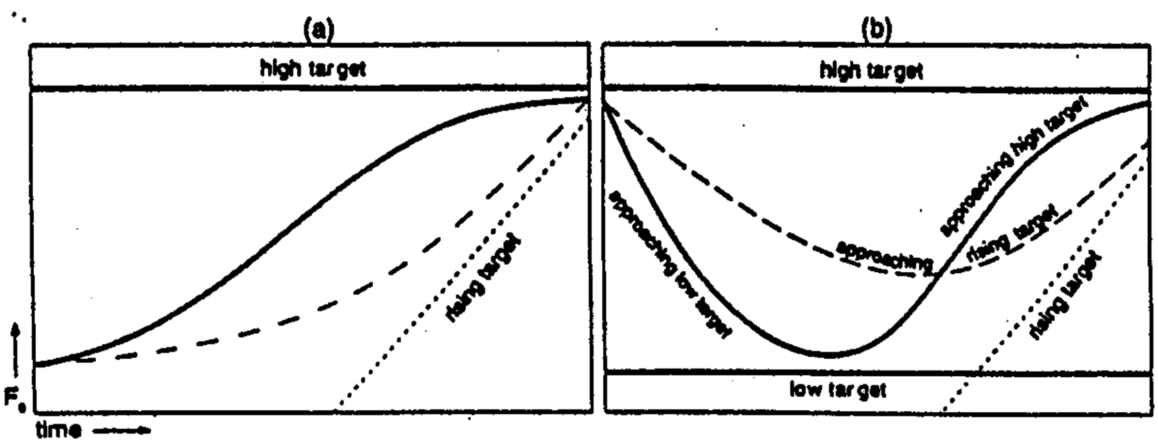


Figure 3: Schematic traces of two different rising tones (a) after low and (b) after high.

Figure 4 shows actual pitch tracks of bi-syllabic inputs. F and R stand for Falling and Rising respectively. In all cases the second tone is rising. Focus on the lower dotted line, which shows a Low+Rise sequence. Xu and Wang observe that this looks very similar to the dashed line in Fig 3 (a) before, suggesting that the target is a rise, not a LH sequence of targets. Now look at the upper thinner solid line, which shows a High+Rise sequence. Again it is a close match to the dashed line in Fig 3(b) earlier, confirming that the target is a rise.

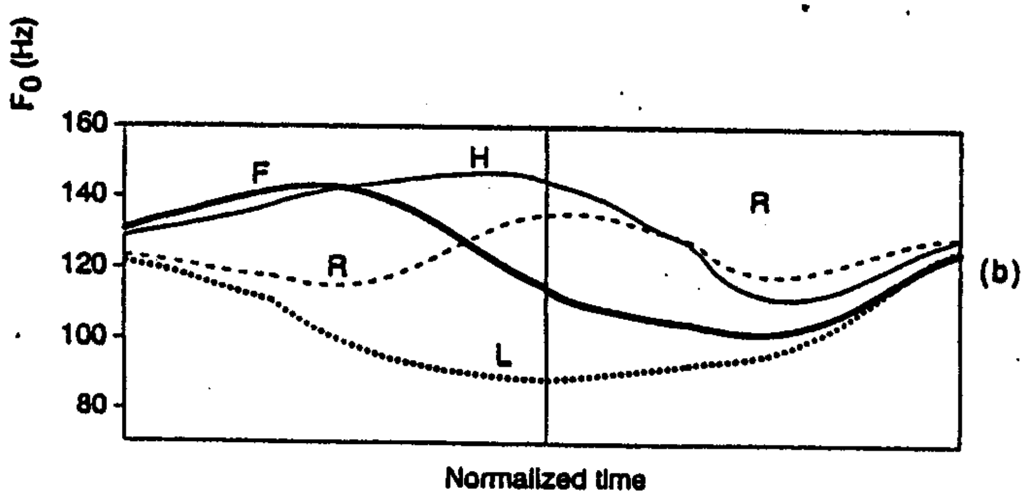


Figure 4: Actual pitch tracks of rising tones after high (H), rising (R), falling (F) and low (L) tones.

The most striking conclusion I draw from these data is that there is no sign whatsoever of a final level target in these pitch tracks. Instead the pitch just keeps rising. This absence of a final level target is supported by the isolation shape of falling tone, which shows no plateau for a putative final L target: Xu: 1997:67. Furthermore the variation in height of this putative target is shown by the traditional transcription as [51] in isolation and [53] medially. Finally, the falling tone does not look like a shorter form of a H.L sequence on two syllables: instead it has a very short high plateau at the beginning, and no low plateau at all at the end.

So what are we to make of these data? To my mind they suggest a radically different view of tonal specification from the prevailing one. I sketch it out here, but it will require extensive further research to determine its full consequences and thus its validity. Suppose that contours are movements *away* from targets. More precisely, suppose that Mandarin high rising and high falling tones are represented as [+Upper, l] and [+Upper, h] respectively, with one target each on the first mora. Level tones have two identical targets, [hh] or [ll], in violation of the OCP, or perhaps one target shared by both moras, as suggested by a reviewer. Further suppose that, at least in Asian languages, register is a syllabic property but tone is a moraic property, and that full-toned syllables are bi-moraic (Duanmu 1990, Tsay 1994). Falling and high syllables will thus have the following characterizations:

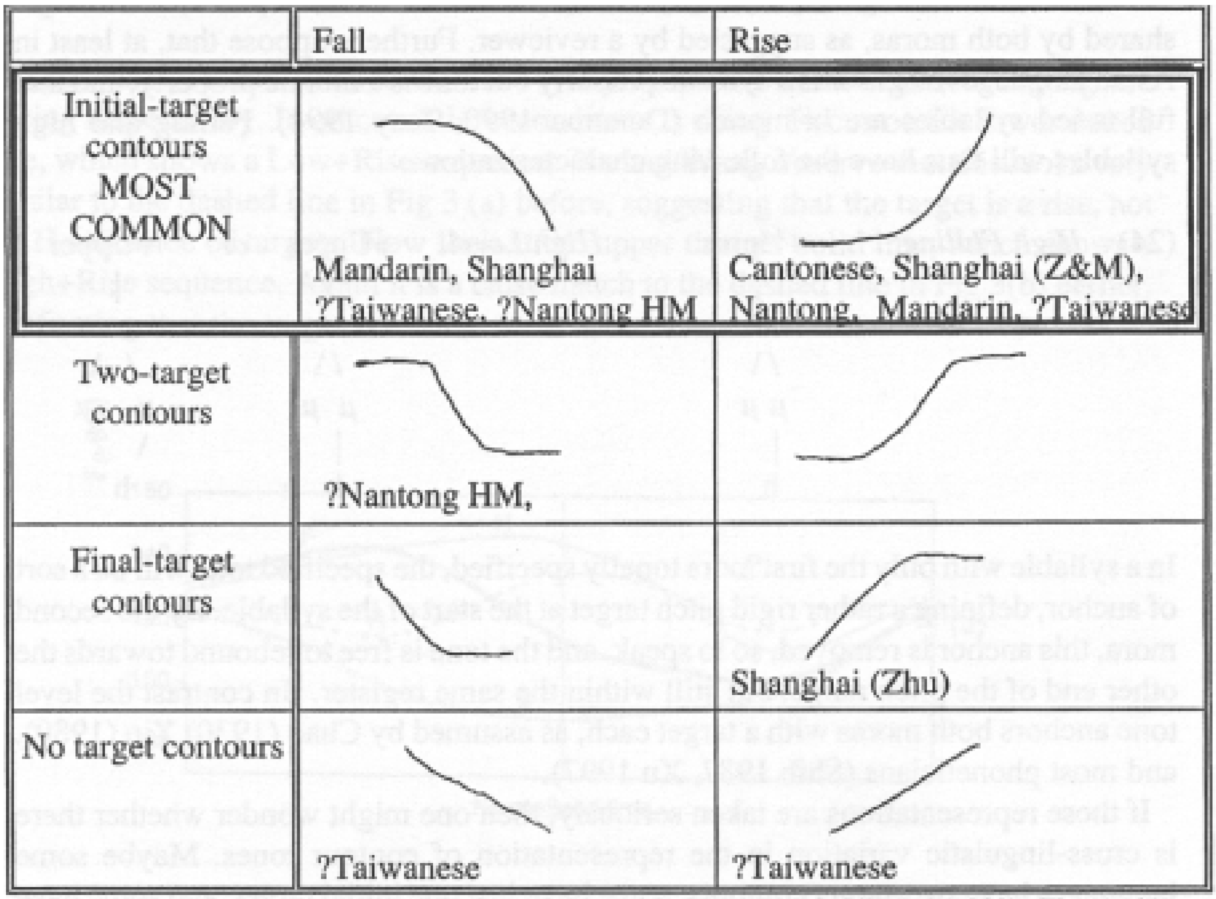
```
(24) High Falling:  +Upper        High Level:   +Upper   or    +Upper
                       |                          |               |
                       σ                          σ               σ
                      / \                        / \             /   \
                     μ   μ                      μ   μ           μ     μ
                     |                          |   |            \   /
                     h                          h   h              h
```

In a syllable with only the first mora tonally specified, the specified tone will be a sort of anchor, defining a rather rigid pitch target at the start of the syllable. By the second mora, this anchor is removed, so to speak, and the tone is free to rebound towards the other end of the pitch range, but still within the same register. [2]In contrast the level tone anchors both moras with a target each, as assumed by Chao (1930) Yip (1980), and most phoneticians (Shih 1987, Xu 1997).

If these representations are taken seriously, then one might wonder whether there is cross-linguistic variation in the representation of contour tones. Maybe some languages have two-target contours, some have just one initial target, and some have just one final target? Interestingly, available phonetic data from Chinese languages overwhelmingly favor the initial target only proposal. The following chart shows schematically what each type of representation would look like phonetically, on the assumption that targets are approached asymptotically (Pierrehumbert and Beckman 1988, Xu 1997). Each cell also includes the names of actual Chinese languages which show these profiles. A question mark shows cases where the sources disagree, or the data is otherwise unclear. For this reason some cases appear in more than one cell of the table. Some cells have no language names. This means that I have found no cases of contours of this shape in my Chinese data, and that I take this to be significant. It remains to be seen whether languages from other areas show a radically different picture. The most common initial-target contours are outlined with a double wavy line.

[2] Since there are contours that start mid and end high or low, such as [35] or [31], the second mora does not necessarily return to neutral pitch. Hence my use of the word “rebound’.

(25) **The predicted phonetic shape of contours under different featural assumptions:**



	Fall	Rise
Initial-target contours MOST COMMON	Mandarin, Shanghai ?Taiwanese, ?Nantong HM	Cantonese, Shanghai (Z&M), Nantong, Mandarin, ?Taiwanese
Two-target contours	?Nantong HM,	
Final-target contours		Shanghai (Zhu)
No target contours	?Taiwanese	?Taiwanese

Nantong (Ao 1993); Mandarin: Xu (1997); Shanghai: Zee and Maddieson (1980), Zhu (n.d.); Cantonese: Chan (1987), Fok (1974) Taiwanese (Peng 1997).

I found only one case that looked like a two-target contour, with plateaux at both ends, and that was some (but not all) tokens of the HM fall in Nantong (Ao 1993). I found one case of a final-target contour: rises in Shanghai as shown by Zhu (n.d.) (but not by Zee and Maddieson 1980). Taiwanese contours (Peng 1997) clearly have no final target, but it is not clear whether they have an initial target or not. All other pitch tracks, from five unrelated dialects, look like initial-target-only contours.

This rather radical initial-target proposal still allows for some partial compositionality of contour tones, but offers a closer fit to the phonetic facts. It makes a number of predictions, whose confirmation or refutation must wait for future research. Firstly, true phonological assimilation/dissimilation should be at the left edge of contours, not the right edge. Second, co-articulation can still be symmetrical, since the phonetic fall or rise caused by the lack of a tonal element can have phonetic effects. This is clearly essential, since progressive co-articulation effects triggered by contour tones are well attested. Xu's Mandarin co-articulation data, I note in passing, which involve progressive assimilation, are clearly phonetic.

The proposal made in this section has far-reaching effects, and it is intended to spark discussion of the issues involved. It is inspired by a sense that phonologists have a duty to take the phonetic interpretation of phonological outputs seriously, and that at present the match between the usual phonological analysis of contours as sequences of levels and their actual $_{F\text{Ø}}$ contours is not close. Even if the specifics of the single-target proposal turn out to be incorrect, this responsibility remains as a challenge.

5.2 Level tones have two targets: Fuzhou

If, as proposed above, contour tones have single tonal targets, it follows that level tones must be distinguished in some way, and the simplest option is to assume that they have two identical targets, resulting in a tonal plateau throughout the syllable. I end with additional phonological evidence in support of this representation for level tones in one dialect.

The data come from Fuzhou (Zheng 1983). Fuzhou full syllables, like Mandarin ones, are bi-moraic feet in their own right, and thus heads. Rhymes of full syllables (first two columns) are minimally [V:], maximally [VGC]. Rhymes of the initial syllables in a particular type of reduplication (third column) are reduced to just V, as in /t^haiŋ/ > [t^ha laiŋ]; the onset change in the second syllable is not relevant here) Full syllables prefer H tones**:** all full syllable inventories (first two columns) have the maximum number of H tones,:{H, HM, MH}. Non-head reduced syllables revert to the usual preference for L, as evidenced by the choice of ML for the contour tones, providing evidence that *H really does dominate *L (third column)**:**

(26) Tone inventories in different positions in Fuzhou (sonorant-final syllables only)

inventory of 5 underlying tones in final syllables	inventory of 4 tones found in ordinary non-final sandhi	inventory of 3 tones found in first-syllable of onset changing sandhi
44 H	44 H	44 H
53 HM	31 ML	31 ML
31 ML	35 MH	
213 LM	53 HM	11 L
242 MH		

Up to this point the Fuzhou analysis closely mimics the Mandarin one, but the Fuzhou data also provide an argument for my earlier proposal that level tones have two targets, not one. Consider the second column. In Fuzhou unlike Mandarin, low level tones are avoided in full syllables, and ML is chosen instead. Under any markedness account, and any ranking of markedness constraints, ML will always lose to plain singleton L, because they tie on *L, and ML also violates *M, and *FALL. The violations incurred by /L/ are a subset of those incurred by /ML/, and thus /ML/ will inevitably lose. For Fuzhou, then, we would expect H, **L**, HM, MH, like Mandarin.

The solution is to treat L as LL, so that lowness is a greater violation the longer it occurs for. This returns to the standard Chinese view, and that of most phoneticians, of level tones as having two targets. The crucial ranking is *L >> *RISE, FALL. This works not only for the sandhi inventory (middle column), but also for the full-toned inventory (first column). Below I show the three possible sets of two tones that could be chosen as optimal to complete a five-tone inventory that has already picked H, HM, MH as three of its members because of high-ranked HEAD=H. It can be seen that the grammar will select LM and ML as the two extra tones, exactly the inventory of the first column above, for final full syllables in Fuzhou.

(27)

	*L	*RISE	*FALL
a. LL, ML	***!		*
b. LL, LM	***!	*	
☞ c. ML, LM	**	*	*

Note that for Mandarin, LL will still be chosen if *RISE, *FALL >>*L.

6. Conclusions

I have made the following proposals. Mid tones are mid because the phonetics realize them at the outer edge of their register, which is by definition the mid-point of the pitch range. Contour tones have an initial target only, whereas level tones have two targets. Dispersion theory explains a number of facts about inventories: (i) Why is M equidistant from H, L? Space-50%**.** (ii) Why do three-tone systems have one of the mid tones missing, and never H or L? Space-50% again**.** (iii) Why are there contour tones at all? To allow for large inventories without spacing level tones too closely. I have left one puzzle completely unresolved: why do Cantonese songs treat two clearly featurally distinct mid tones the same? This unsolved case was left in on purpose to induce humility: there is an enormous amount about tones that we still do not understand, and that hopefully will spur further research.

Selected References

Anderson S. 1978. Tone features. In V. Fromkin, ed.,*Tone: A linguistic survey*. Academic Press. New York. 133-176.

Ao, Benjamin. 1993. *Phonetics and phonology of Nantong Chinese.* PhD Dissertation, The Ohio State University.

Bao, Zhi-ming. 1990 .*On the nature of tone.* Ph D Dissertation. MIT

Bao, Zhi-ming. 1999 The structure of tone. Oxford University Press.

Boersma, Paul 1997 *The elements of functional phonology*. PhD Dissertation, University of Amsterdam. ROA#-173

Chan, Marjorie 1987 Tone and melody interaction in Cantonese and Mandarin Songs. UCLA WPP 68: pp132-169

Chao, Yuan-ren 1930 A system of tone letters. Le Mâitre Phonetique. 45:24-47

Chen, Matthew 2000 *Tone Sandhi*. Cambridge University Press.

Cheng, C.C. 1973. A quantitative study of Chinese tones. Journal of Chinese Linguistics 1: 93_110.

Clumeck, H. 1980 The acquisition of tone. In George D. Allen and Sarah Hawkins, eds. *Child Phonology Vol: Production* : 257-275.

de Lacy, Paul 1998 Tone and Prominence. Ms, UMass, Amherst.

Duanmu, San. 1990. *A Formal Study of Syllable, Tone, Stress and Domain in Chinese Languages.* Ph.D. Dissertation, MIT.

Fabb, N. 1997 *Linguistics and literature.* Blackwell, Oxford.

Flemming, Edward 1995 Auditory representations in phonology. PhD Dissertation, UCLA.

Fok, Chan Yuen-yuen 1974 A perceptual study of tones in Cantonese. Center of Asian Studies. University of Hong Kong.

Hashimoto, Anne Oi-Kan Yue. 1980. Tone sandhi across Chinese dialects. In *Wang Li Memorial Volumes, English volume*, edited by The Chinese Language Society of Hong Kong. Joint Publishing Co (HK). pp445-474.

Hayes, B. 1996. Phonetically driven phonology: the role of Optimality Theory and inductive grounding. *Proceedings of the 1996 Milwaukee Conference on Formalism and Functionalism in Linguistics.* ROA # -158

Hsieh, Hsin-I 1976 On the unreality of some phonological rules. *Lingua* 38:1-19.

Hyman, Larry. 1993. Register tones and tonal geometry. In K. Snider and H. van der Hulst (eds.), The Phonology of Tone: The Representation of Tonal Register (pp. 75_108). Berlin: Mouton de Gruyter.

Jun, Jongho 1995 *Perceptual and articulatory factors in place assimilation: an Optimality Theoretic approach.* PhD Dissertation, UCLA

Kirchner, Robert 1996 Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27.2: 341-349

Kirchner, Robert. 1998. An effort-based approach to consonant lenition. PhD Dissertation, UCLA.

Li, C. and S. Thompson. 1978. The acquisition of tone. In V. Fromkin ed. Tone; A Linguistic Survey. Academic Press, London. pp 271-284.

Lindblom, B. 1990 Phonetic content in phonology. *PERILUS* 11.

Mohanan, K. P. 1986. The Theory of Lexical Phonology. Dordrecht: Reidel.

Padgett, Jaye. 1997. Surface contrast in phonology. Talk given at UC Irvine, April 1997.

Peng, Shu-hui 1997 Production and perception of Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics* 25: 371-400

Pierrehumbert, J. and M. Beckman 1988 *Japanese tone structure.* MIT Press. Cambridge.

Pulleyblank, Douglas. 1986. Tone in Lexical Phonology. Dordrecht: D. Reidel.

Shih, Chi-lin. 1987. The phonetics of the Chinese tonal system. Ms., AT&T Bell Labs Silverman, D. 1995. *Phasing and recoverability.* PhD Dissertation, UCLA.

Steriade, D. 1995 Neutralization and the expression of contrast. Ms., UCLA.

Sundberg, J. 1979 Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics* 7:71-79.

Tsay, J. 1994. *Phonological pitch.* Ph.D. Dissertation. University of Arizona.

Tsay, J. and J. Myers 1996 Taiwanese tone sandhi as allomorph selection. *BLS* 22: 394-405.

Wright, M. 1983 A Metrical Approach to Tone Sandhi in Chinese Dialects". Ph.D. Dissertation, U. of Massachusetts, Amherst

Xu, Yi 1997 Contextual tonal variations in Mandarin. *Journal of Phonetics* 25: 61-83

Xu, Yi and Q. Emily Wang 1999 Pitch targets and their realization. To appear in *Speech Communication*.

Yip, M. 1980. *The Tonal Phonology of Chinese*. MIT PhD Dissertation. Published 1990 Garland, New York.

Yip, M. 1989 Contour tones. *Phonology* 6.1 :149-174.
Zee, E. and I. Maddieson 1980 Tones and tone sandhi in Shanghai: phonetic evidence and phonological analysis. *Glossa* 14.1:45-88.
Zhang, Jie 1999. Phonetic duration effects on contour tone distribution. Ms, UCLA
Zhang, Sheng Yu. 1979. Chaoyang fangyan de chongdieshi. [Reduplication in the Chaoyang dialect]. *Zhongguo Yuwen* 1979, 2.106-114.
Zhang, Sheng Yu. 1981. Chaoyang fangyan de yuyin xitong. [An outline of Chaoyang phonology]. *Fangyan* 1981, 1.27-39.
Zhang, Sheng Yu. 1982. Chaoyang fangyan de xiangshengzi chongdie shi. [Reduplicated onomatopoeic particles in the Chaoyang dialect]. *Fangyan* 1982, 3.181-182.
Zhu, Xiaonong (n.d.) Stress patterns of Shanghai disyllabic words: Acoustic correlates and phonological implications. Ms, Hong Kong Polytechnic University.
Zoll, Cheryl. 1997. A note on multiple prominence and tone mapping. *MIT Working papers in Linguistics.* 30: 97-111.