LANGUAGE AND LINGUISTICS 5.4:925-968, 2004
2004-0-005-004-000099-1

Contour Tone Licensing and Contour Tone Representation

Jie Zhang
University of Kansas

In this paper, I argue that the positional restrictions of contour tones do not result from their bimoraicity requirement for licensing. The arguments come from the licensing advantage of syllables in positions that are not moraically privileged, the extra levels of markedness distinction in contour tones that go beyond the maximum mora count allowed in the theory, the differences among contour tones with the same number of pitch targets, and the long lasting problem of moraic inconsistency. I propose a theoretical apparatus that allows more phonetic detail in both the tonal shape and rime duration to enter into phonological representations and show that it on the one hand provides a better account for the cross-linguistic behavior of contour tone licensing, but on the other hand does not endanger the predictive power of phonological theory.

Key words: contour tone, mora, duration, phonology-phonetics interface

1. Introduction

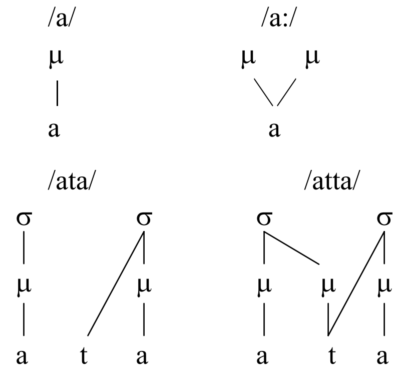
1.1 The general roles of the mora in phonology

The notion of the mora in linguistic theory can be traced back to Trubetzkoy (1939), in which he acknowledged its role in the placement of stress in Classical Latin: '[It] always occurred on the penultimate "mora" before the last syllable, that is, either on the penultimate syllable, if the latter was long, or on the antepenultimate, if the penultimate was short.' (Trubetzkoy 1939, Baltaxe translation 1969:174) It was then referred to in McCawley's (1968) study of Japanese accent to account for the occurrence of different pitches on a single rime and Newman's (1972) survey of stress assignment in languages in which the distinction between heavy and light syllables must be made. It entered into generative phonology as a level of representation starting from Hyman's (1985) proposal of the weight unit *x*, which was equivalent to the mora. McCarthy & Prince (1986) and Hayes (1989) explicitly proposed the mora tier in the representation and argued that the moraic representation was what motivated weight-related phenomena such as stress assignment, tone bearing, and compensatory lengthening. For an overview of the history and arguments for the mora, see Broselow (1995).

In essence, the mora plays two roles in phonology. First, it can be used to represent segment length: a vowel length contrast can be expressed through a mono- vs. bimoraic distinction as in (1a), while singleton and geminate consonants may differ in that the former is non-moraic while the latter is monomoraic, as in (1b) (McCarthy & Prince 1986, Hayes 1989).

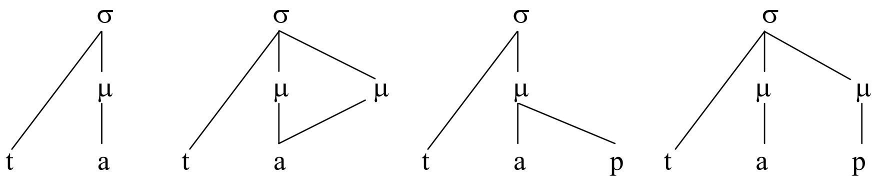
(1) a. /V/ vs. /V:/

b. /C/ vs. /C:/



More importantly, the mora can be used to represent the heavy vs. light distinction in weight-related processes. A heavy syllable has two moras while a light syllable has one. Hayes (1989) proposes that CV is inherently monomoraic and CVV bimoraic; for CVC, a language has the choice of treating the coda as moraic or not. The moraic representations for CV, CVV, and CVC are given in (2).

(2) a. CV b. CVV c. CVC (light) d. CVC (heavy)



Consequently, the mora encodes the asymmetries between onsets and codas in weight-related processes. E.g., in stress assignment, the presence of an onset never determines the stressability of the syllable (but see Everett & Everett 1984), while the presence of a coda often does; in compensatory lengthening, the loss of a coda triggers lengthening of the nucleus, while the loss of an onset rarely does (Hayes 1989); in templatic morphology, the onset of a syllable template is often optional, while the coda rarely is (Broselow 1995). The way in which these asymmetries are expressed in moraic theory is that onsets are never mora-bearing, while codas may be mora-bearing through language-specific rules. One major advantage of the moraic theory is that all weight-related phenomena in a language are now subsumed under a single moraic representation.

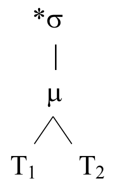
1.2 Moraic licensing of contour tones

The mora has also been proposed to be the Tone-Bearing Unit, most notably by Duanmu (1990, 1994a, b), who made the observation that in Chinese languages, the distributional restrictions of contour tones in a particular language are often correlated with its inventory of rime types and rime duration. Specifically, he argues that Chinese languages fall under two major categories regarding syllable rimes: one has coda contrasts, diphthongs, and longer rime duration, such as Mandarin, Cantonese, and Fuzhou; the other has no coda contrasts, no diphthongs, and shorter rime duration, such as Shanghai, Suzhou, and Danyang. Crucially, in the first type of languages, distributional restrictions on contour tones are few and far between, while in the second type, tone sandhi processes that turn contour tones into levels abound. These observations are summarized in (3).

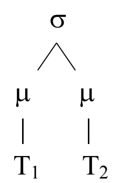
(3)	Coda contrast	Diphthongs	Rime duration	Contour distribution
Mandarin	Yes	Yes	Long	Freely
Cantonese	Yes	Yes	Long	Freely
Fuzhou	Yes	Yes	Long	Freely
Shanghai	No	No	Short	→ Level
Suzhou	No	No	Short	→ Level
Danyang	No	No	Short	→ Level

To capture the relation between rime properties and contour tone licensing, Duanmu argues that a contour tone must be represented as a concatenation of level tones, each of which needs a mora to be licensed, and that the difference in contour restriction between the two types of languages stems from the fact that syllables are bimoraic in the former and monomoraic in the latter, as indicated in (4).

(4) a. No contour tone on monomoraic syllables



b. Contour tone allowed on bimoraic syllables



He argues against Yip's (1989) proposal that contour tones in Chinese are phonological units and properties of the syllable, on grounds that the correlation between rime duration and contour restrictions cannot be captured in such an approach.

In this paper, I argue that a careful review of the contour tone licensing behavior in Chinese and elsewhere suggests that bimoraicity is neither a sufficient nor a necessary licensing condition for contour tones, and that contour tone licensing can be better understood if more phonetic detail in both the tonal shape and rime duration are allowed to enter into phonological representations.

2. Against the mora

2.1 Advantage of prosodic-final syllables in contour tone licensing

The first argument against moraic licensing comes from the fact that contour tones are not only favored on syllables with indisputably higher mora count, such as those with phonemically long vowels or sonorant codas. As noted by Clark (1983), many African languages license contour tones more freely in prosodic-final positions. E.g., in Kikuyu, when a floating High docks onto a syllable with an original Low, if the syllable is not phrase-final, the High replaces the Low; but if it is, a $\widehat{\text{LH}}$ Rise results, as shown in (5).

(5) Kikuyu High tone docking:

kàriòkì	moèɣà	→	kàriòkí	moèɣà	'good Kiriuki'
L H	L		L H	L	
kàriòkì		→	kàriòkǐ		'Kiriuki'[1]
L H			L H		

Clark's intuition was corroborated in a survey of contour tone distribution that I conducted (Zhang 2002). Of the 187 languages surveyed, forty-seven show the preference for prosodic-final syllables as contour bearers, as shown in (6).

[1] Clements & Ford (1979) argue that there is also a floating extra Low after the floating H for the word. The extra Low is not realized in either of these forms.

(6) Advantage of prosodic-final position for contour tone bearing:

Language phylum	Number of languages	Languages
Afro-Asiatic	5	Agaw (Awiya), Bolanci (Bole), Galla (Booran Oromo), Rendille, Sayanci
Daic	1	Ron Phibun Thai
Niger-Congo	19	Bandi, Kivunjo Chaga, Machame Chaga, Etung, Gã, Kenyang, Kikuyu, Kisi, Kɔnni, Nana Kru, Wobe Kru, Kukuya (Southern Teke), Lama, Luganda, Mende, Ngamambo, Ngazija, Ngie, Ngumbi, Tiv
Nilo-Saharan	2	Bari, Lulubo
Oto-Manguean	3	San Andrés Chichahuaxtla Trique, San Juan Copala Trique, Mitla Zapotec
Sino-Tibetan	15	Beijing, Fuzhou, Kunming, Nanchang, Nanjing, Pingyang, Pingyao, Shuozhou, Suzhou, Shexian, Wuhan, Wuyi, Xining, Xinzhou, Yanggu
Trans-New Guinea	1	Mianmin

As we can see, the preference to have contour tones on prosodic-final syllables is indeed widely attested in African languages. E.g., in Etung, Emondson & Bendor-Samuel (1966) observe that while there is no restriction on the occurrence of level tones (H, L, $^{!}$H), the falling and rising contours (H͡L, H͡$^{!}$H, L͡H) are restricted to the final syllable of mono- and disyllabic phonological words.[2] In Luganda, there is a vowel length contrast in open syllables. But this contrast is neutralized to [-long] in final position. The tonal inventory includes H, L, and H͡L, with no positional restriction on H and L, but H͡L is disallowed on CV except in final position (Snoxall 1967, Stevick 1969, Hyman & Katamba 1990, 1993).

[2] Edmondson & Bendor-Samuel account for the distribution by identifying 12 patterns of tonal melody (LLL, LHH, $^{!}$HHH, HLL, HHH, LHL, LLH, HHL, H$^{!}$HH, HLH, LH$^{!}$H, HH$^{!}$H) and mapping them left-to-right onto syllables in a phonological word. But as Morris Halle pointed out (Goldsmith 1976, p.156, fn.2), the eight patterns that do not involve $^{!}$H include all the melodies expressible as $T_1T_2T_3$. This observation makes the melodic approach immediately less attractive here, as it does not serve the purpose of restricting possible tonal combinations in polysyllabic word. Furthermore, Zhang (2000a) argues that even if tonal melodies can be independently motivated, the effect of left-to-right mapping cannot be rendered in OT without specifically referring to the final position as privileged for contour tone bearing.

The final preference for contour tone licensing is also attested in many other language phyla. Chinese dialects in particular provide many instances of the manifestation. Beijing Chinese represents a typical scenario: the convex tone 213 is only realized as such word-finally, and is simplified to a simple contour 21 or 35 when preceding another syllable, as shown in (7).

(7) Beijing tone sandhi:

a.	55-**213** → 55-**213**	b.	**213**-55 → **21**-55
	35-**213** → 35-**213**		**213**-35 → **21**-35
	213-**213** → 35-**213**		**213**-213 → **35**-213
	51-**213** → 51-**213**		**213**-51 → **21**-51

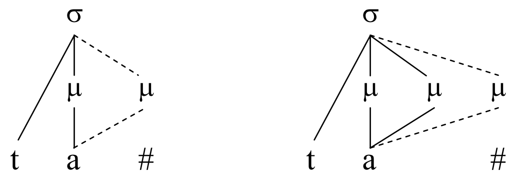
Phonetically, the privileged status of prosodic-final position for contour tone bearing is of no surprise, since prosodic-final lengthening is well attested in many languages; and contour tones, which involve articulatory movement and perceptual decoding of pitch change, demand sufficient duration in their realization. Moreover, Zhang (2002) documents that in Beijing Chinese, word-final syllables have considerably longer rime duration than non-final syllables (213ms vs. 151ms; 2 speakers, 8 pairs each); in Luganda, neutralized final short vowels are considerably longer than non-final short vowels (final CV: 179ms; non-final CV: 117ms; CVV: 293ms; 1 speaker, 12 tokens each for final and non-final CV, 7 tokens for CVV).

If bimoraicity was the licensing condition for contour tones, we would expect the privileged final position to have moraic advantage over non-final positions. But for the following reasons, the durational advantage of final position cannot always be appropriately captured moraically.

First, although there are languages that neutralize the vowel length contrast in final position, such as Luganda, final lengthening is by no means always neutralizing. But the final advantage for contour tones is not restricted to languages that have neutralizing final lengthening. If the mora count is restricted to representing structural differences, then it is not an appropriate measure for non-neutralizing final-lengthening and hence cannot be a unified explanation for the gravitation of contour tones to final position. Moreover, even when the final position is length-neutralizing, it is often to [-long], not [+long], as in Luganda. Thus there is in fact motivation for subtracting moras, not adding them, in final position.

One could argue that the motivation for introducing moras is not necessarily structural, therefore the addition of a mora to the final syllable even when length is not neutralized is allowed, as in (8).

(8) a. CV# b. CVV#



But the mora introduced here is apparently for the purpose of contour tone bearing alone and is typologically at odds with systems of weight-sensitive stress. Hayes (1995) shows that there are very few cases in which the final syllable is guaranteed stress regardless whether it is heavy or light, while non-final syllables are only stressed when they are heavy;[3] But cases in which the final syllable is at a *disadvantage* for attracting stress abound: English, Estonian, Arabic dialects, Spanish, Romanian, Ancient Greek, Menomini, etc. This fact is in direct conflict with the proposal above if we assume that the moraic structure is the basis for all weight-related processes.

2.2 Advantage of syllables in shorter words in contour tone licensing

The survey of contour tone distribution also reveals that syllables in shorter words are better contour tone licensers than syllables in longer words. Of the 187 languages, nineteen show this asymmetry, as shown in (9).

(9) Advantage of syllables in shorter words for contour tone bearing:

Language phylum	Number of languages	Languages
Afro-Asiatic	4	Galla (Booran Oromo), Margi, Musey, Rendille
Niger-Congo	7	Abidji, Etung, Gã, Kinyarwanda, Kukuya (Southern Teke), Mende, Ngamambo
Sino-Tibetan	7	Changzhou, Chaoyang, Chengdu, Chongming, Lüsi, Ningbo, Shanghai
Trans-New Guinea	1	Siane

Seven languages that exhibit this effect are Chinese dialects, and among them, five

3 Tübatulabal, Aklan, and Cebuano are the only cases cited by Hayes that display this behavior. E.g., in Tübatulabal, final syllables and heavy syllables (CV:) are stressed, and every other light syllable (CV) before a heavy syllable is stressed.

(Changzhou, Chongming, Lüsi, Ningbo, and Shanghai) are Wu dialects. The tone sandhi of the Northern Wu dialects is typically described as “left-dominant”, while that of the Southern Wu dialects, “right-dominant” (Yue-Hashimoto 1987). In these dialects, the tone of the “dominant” syllable in a polysyllabic word often spreads across the whole word, taking over the lexical tones of other syllables. E.g., the schematic in (10) encapsulates the data pattern in Shanghai, a Northern Wu dialect: the pitch targets on the initial syllable of a polysyllabic compound are realized throughout the compound, and consequently, even if monosyllabic morphemes have contour tones in isolation, no contour surfaces in polysyllabic compounds (Zee & Maddieson 1979, Duanmu 1990).[4]

```
(10) a.  σ σ σ σ     σ σ σ σ     b.  σ σ σ σ     σ σ σ σ
          /\      →  | | | |          /\      →  | | | |
         H  L        H L L L         L  H        L H L L
```

When complex contour tones with three pitch targets are present in monosyllabic morphemes, spreading mechanisms similar to (10) will ensure that no such complex contours will surface in disyllabic words. E.g., in Changzhou (Wang 1988), a disyllabic word with 523 followed by any tone will be realized as 5-23 tonally.

The two non-Wu Chinese dialects here—Chaoyang (Zhang 1979, 1980) and Chengdu (Cui 1997)—both have a complex contour as one of the lexical tones (213 in Chengdu and 313 in Chaoyang). But in both dialects, the complex contour only surfaces in monosyllabic citation forms. In disyllabic forms, tone sandhi occurs and the complex tone is simplified. In Chaoyang, 313 surfaces as 33 when occurring on the first syllable, and as 11 when occurring on the second syllable. In Chengdu, 213 is realized as 13 in any disyllabic or polysyllabic utterances. We might not be able to predict from phonetics the exact shape of the sandhi tones in these dialects, but at least we are able to restrict the inventories from which the sandhi tones are drawn.

Mende, a Mande language spoken in Sierra Leone, is another example. Zhang (2000a) summarizes the contour tone restrictions in Mende as in (11). In particular, a long vowel can carry a $\mathrm{L\widehat{H}L}$ complex contour in monosyllabic words, but not in longer words, and a short vowel can carry a $\widehat{\mathrm{LH}}$ contour in monosyllabic words, but not in longer words. Mende also exhibits the prosodic-final advantage: a short vowel can carry a $\widehat{\mathrm{HL}}$ contour word-finally, but not elsewhere.

4 Note again that the position taken here is that the mechanism of “tonal melody spreading” does not constitute an *explanation* for the contour tone restriction. The reason that such processes take place is an articulatorily and perceptually well-motivated one, not a purely formal one. Otherwise, we would be left without an explanation for why processes that *create* contour tones on syllables in longer words, which are not hard to conceive formally, are such a rarity.

(11) Mende contour tone restrictions:

Vowel length	Syllable count in word	Syllable position in word	$\widehat{\text{LHL}}$ ok?	$\widehat{\text{LH}}$ ok?	$\widehat{\text{HL}}$ ok?
VV	1	final	**yes**	**yes**	**yes**
VV	>1	any	no	**yes**	**yes**
V	1	final	no	**yes**	**yes**
V	>1	final	no	no	**yes**
V	>1	non-final	no	no	no

Phonetic studies on English and Swedish have established, *ceteris paribus*, that syllables in shorter words are longer than syllables in longer words (Lehiste 1972, Lindblom et al. 1981, Lyberg 1977, Strangert 1985). More relevantly, in Shanghai, Zee & Maddieson (1979) show that the syllable nucleus in a monosyllabic domain has an average duration of 327ms, while Duanmu (1994b) shows that the syllable nucleus in a polysyllabic domain is considerably shorter at 162ms.

One may again assume that syllables in shorter words simply have more moras. This is not immediately unreasonable, as word minimality sometimes does require monosyllabic words to have a long vowel. But again, the lengthening effect in short words is not always neutralizing, as in Mende. Then the use of moraic licensing here is open to the same criticism as using it to capture the prosodic-final preference for contour tones.

2.3 Levels of distinction in contour tone licensing ability

Given that the primary role of the mora is to capture the heavy vs. light or long vs. short distinction, the maximum mora count of a syllable is usually assumed to be two. This is the position taken by McCarthy & Prince (1986) and Steriade (1991). Hayes (1989) argues that sometimes three levels of weight or length distinction need to be made. E.g., in Estonian, there is a three-way length contrast for vowels (Harms 1962, Tauli 1973); in a dialect of Hindi, superheavy syllables (CVVC, CVCC) behave like a heavy syllable followed by a light syllable; in Persian metrics, superheavy (CVVC, CVCC) and ultraheavy (CVVCC) syllables are scanned as a long position followed by a short position /_ ‿/ (Elwell-Sutton 1976, Hayes 1979). But to the best of my knowledge, no claim has been made that more than three levels of weight or length distinction are necessary. As an illustration, the Persian example above shows that an ultraheavy syllable does not have a different metrical scansion from the trimoraic superheavy syllables. Contour tone licensing, however, sometimes requires more than three levels of distinction. E.g., the Mende case we have seen in (11) requires four: syllables that can carry L͡HL, syllables that can carry L͡H, syllables that can carry H͡L, and syllables that cannot carry any contour. This goes beyond the maximum mora count that any version of moraic theory is willing to accommodate.

2.4 Differences among tones with the same number of pitch targets

2.4.1 The fall vs. rise asymmetry

One other distributional property worth noting in Mende is that the rising tone has a more restricted occurrence than the falling tone—the fall can occur on a short vowel in the final position of a polysyllabic word, while the rise cannot. The fall-rise asymmetry in Mende is not an isolated incident, as it is reflected in multiple ways in the survey in Zhang (2002).

First, thirty-seven languages in the survey do not allow surface rise, while only three do not allow surface fall. One reason that leads to this asymmetry is that when rising tones arise through spreading or morphological concatenation, they tend to be simplified. E.g., in Lama, when a Low spreads rightward to a final syllable with a High, the Rise that would have surfaced is realized as Downstep; but when a High spreads rightward to a final syllable with a Low, the H͡L surfaces intact, as in (12) (Ourso 1989).

(12) Lama rising simplification:

na: tẹ → ná: tꜜẹ 'chez cow' wa:l tẹ → wá:l tệ 'under husband'

H L H H L

Second, rising tones are often more restricted in distribution than falling tones. In Kɔnni (Cahill 1999) and Tiv (Pulleyblank 1986), contour tones are restricted to word-final position, but the Rise is further restricted to CVR and CVVR (R=Sonorant), while the Fall is allowed on CV.

Third, rising tones, when realized, tend to lengthen the rime. E.g., in Gã, there is a vowel length contrast; on a phrase-final short vowel, the Fall can occur without lengthening, but the Rise can only occur upon neutralizing lengthening, as in (13) (Paster 1999). In Mitla Zapotec, there is no vowel length contrast; the rise induces non-neutralizing lengthening, while the Fall does not (Briggs 1961).

(13) Gã neutralizing lengthening:

he → hê 'to buy' cha → chàá 'dig!'

H L L H (imperative)

The fall-rise asymmetry is grounded in articulation, as Sundberg (1979) has shown that a pitch rise takes longer to implement than a pitch fall with equal excursion. But this asymmetry is difficult to represent moraically, as the same number of pitch targets ought to require the same number of moras to be licensed, as shown in (14).

(14) a.

```
    σ
   / \
  μ   μ
  |   |
  L   H
```

b.

```
    σ
   / \
  μ   μ
  |   |
  H   L
```

2.4.2 The effects of excursion size

The distributional differences between contour tones with different magnitudes of pitch excursion are another aspect of the same problem, and they are widely manifested in the different contour tone licensing abilities between checked (syllables closed with an obstruent /p/, /t/, /k/, or /ʔ/) and non-checked syllables (open syllables or syllables closed with a nasal /m/, /n/, or /ŋ/) in many Chinese dialects (Pingyao, Fuzhou, Shanghai, Cantonese, etc.) (Zhang 1998). E.g., in Pingyao Chinese, the tonal inventory includes three contour tones 13, 35, and 53. All three tones can occur faithfully on non-checked syllables; but on checked syllables, 35 does not occur, and 13 and 53 are partially flattened to 23 and 54, as illustrated by the following examples (Hou 1980).

(15) Pingyao examples:

13	pu	'to hatch'	iŋ	'overcast'
23	pʌʔ	'to push aside'	xuʌʔ	'hair'
35	pu	'cloth'	tuŋ	'to move'
53	pu	'to mend'	tiŋ	'nap'
54	pʌʔ	'a musical instrument'	xuʌʔ	'to live'

The phonological affinity between 13, 53 on the one hand and 23, 54 on the other is not only manifested by their phonetic similarities, but also by the identity of their phonological behavior. Tone sandhi behavior in Pingyao is syntactically conditioned. Words in different syntactic configurations have different sandhi forms even if they have the same base form. Hou (1980) summarizes the tone sandhi behavior of disyllabic words of predicate-object or subject-predicate configuration as in (16). The leftmost column and the top row show the base forms of the first and second syllables respectively. The body of the table indicates the sandhi forms of the disyllabic words. Checked tones are underlined for easy identification.

(16)

$\sigma_1 \backslash \sigma_2$	13	$\underline{23}$	35	53	$\underline{54}$
13	13-13	13-$\underline{23}$	31-35	35-423	35-$\underline{423}$
$\underline{23}$	$\underline{23}$-13	$\underline{23}$-$\underline{23}$	$\underline{32}$-35	$\underline{45}$-423	$\underline{45}$-$\underline{423}$
35	13-13	13-$\underline{23}$	31-35	35-423	35-$\underline{423}$
53	53-13	53-$\underline{23}$	53-35	35-423	35-$\underline{423}$
$\underline{54}$	$\underline{54}$-13	$\underline{54}$-$\underline{23}$	$\underline{54}$-35	$\underline{45}$-423	$\underline{45}$-$\underline{423}$

Tone sandhi behavior for disyllabic words with structures other than predicate-object or subject-predicate, such as modifier-noun, verb-verb, noun-noun, and predicate-adjunct, is given in (17).

(17)

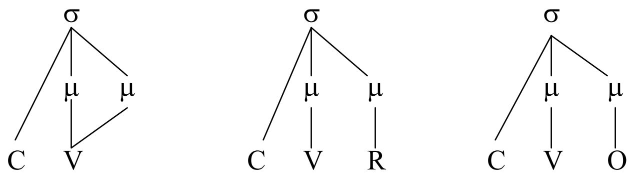
$\sigma_1 \backslash \sigma_2$	13	$\underline{23}$	35	53	$\underline{54}$
13	31-35	31-$\underline{45}$	13-13	31-53	31-$\underline{54}$
$\underline{23}$	$\underline{32}$-35	$\underline{32}$-$\underline{45}$	$\underline{23}$-13	$\underline{32}$-53	$\underline{32}$-$\underline{54}$
35	35-53	35-$\underline{54}$	35-53	35-53	35-$\underline{54}$
53	53-13	53-$\underline{23}$	53-35	**53-53**	53-$\underline{54}$
$\underline{54}$	$\underline{54}$-13	$\underline{54}$-$\underline{23}$	$\underline{54}$-35	**$\underline{45}$-53**	$\underline{54}$-$\underline{54}$

For an account of the tone sandhi behavior, see Zhang (1999). For the present purpose, it is sufficient to note that in both types of sandhi, 13 and 23 have the same behavior, so do 53 and 54, except the pair in boldface in (17), which I take as an exception.

The restriction of pronounced contours to longer rimes has a clear articulatory basis, as the implementation of farther-apart pitch targets would require the relevant laryngeal muscles to contract or relax to a greater degree, and thus demand a greater duration of its carrier. But the relevance of duration is again hard to capture moraically, as $\widehat{\mathrm{LM}}$ also requires two moras to be licensed, just as $\widehat{\mathrm{LH}}$ (cf. (14)).

Duanmu (1990, 1994a) proposes that in isolation, syllables in Chinese dialects are generally bimoraic: the vowel in CV is lengthened; the coda consonant, whether it is a sonorant or an obstruent, always contributes a mora to the syllable. The usual lack of contour tones on CVO syllables is due to low-level phonetic reasons: since the obstruent coda in Chinese is usually unreleased, a tone cannot be phonetically realized on it, even though it may be underlyingly linked to the mora that the coda contributes. The proposed moraic representations for CVV, CVR, and CVO are shown in (18).

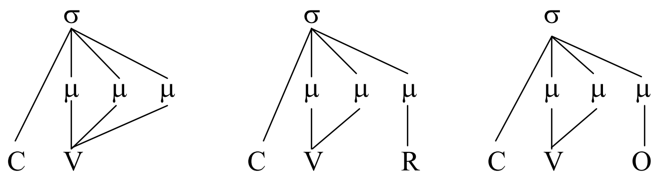
(18) a. CVV b. CVR c. CVO



In languages like Pingyao, which allow contour tones on CVO, Duanmu argues that the vowel in CVO is lengthened to bimoraic. This allows the two level tones that comprise the contour tone to be both realized phonetically. But this essentially leaves the smaller pitch excursion of the contour tones on CVO unaccounted for. Duanmu (p.c.) has suggested two possible solutions.

First, the vowels in CVV and CVR may also be lengthened, which will render all syllable types trimoraic, as shown in (19). But the problem then again becomes why complex contours do not occur on CVV and CVR in this language: there is no reason why the lengthening of the vowel in CVV and CVR should not license one more pitch target as the extra mora in CVO does.

(19) a. CVV b. CVR c. CVO



Second, the pitch excursion reduction is a phonetic effect, i.e., it falls outside the purview of phonology. Even though the vowel in CVO is bimoraic, it is phonetically shorter than the bimoraic vowel in open syllables or the vowel+sonorant combination in CVR. This phonetic shortening gives rise to a phonetic contour flattening effect on CVO. Yip (1995), though she disagrees with Duanmu's view that the mora is the tone-bearing unit and that there is no contour tone unit, endorses the phonetic nature of partial contour flattening. I argue against this position in the following section.

2.4.3 Partial contour flattening is not outside the purview of phonology

Aside from the fact that it cannot be extended to the rise-fall asymmetry, the proposal to treat partial contour reduction as low-level phonetic effects is fundamentally flawed in the following ways.

First, when a pronounced pitch excursion encounters a carrier with short duration, different languages adopt different strategies to resolve the conflict. Among them are: neutralizing contour flattening as in Xhosa, which reduces a Fall to level High in unstressed positions (Lanham 1958, 1963); partial contour flattening as in Pingyao Chinese; neutralizing rime lengthening as in Gã; partial rime lengthening as in Mitla Zapotec; and a combination of partial flattening and lengthening as in Hausa, which allows the occurrence of a falling tone on CVO (O=obstruent) upon simultaneously reducing the pitch excursion of the fall (as compared to the fall on CVV and CVR) and lengthening of the vowel (Gordon 1999, Zhang 2002). This suggests that, as one of many possible strategies, partial flattening does have linguistic relevance and cannot be simply regarded as phonetic. It would be desirable to have a unified theory that accounts for all the strategies that languages employ, especially that there is a clearly identifiable phonetic parameter that all the contour licensing conditions relate to—sonorous rime duration.

Second, from the Pingyao data alone, it is conceivable to consider 23 and 54 to be incomplete phonetic realizations of 13 and 53 on a short duration. But there are many languages, especially in Sino-Tibetan, in which the tones on CVO generally have smaller pitch excursions than those on CVV and CVR, but there is no clear resemblance between the two sets of tones in either phonetic similarity or sandhi behavior.

For example, in Xiamen (Chen 2000), a Southern Min dialect of Chinese, five tones can occur on CV and CVR syllables—44, 24, 53, 21, and 22, and two tones can occur on CVO syllables—32 and 4. It is not immediately obvious whether the small fall 32 on CVO is a natural phonetic reduction of any of the tones on CVV and CVR. Moreover, if we look at the tone sandhi behavior of Xiamen, we can see that 32 behaves quite differently from the tones on CV and CVR. Xiamen tone sandhi is sensitive to prosodic context, but not to tonal context. So each tone in non-phrase-final position is changed into another tone regardless of the tone following it, as schematically shown in (20). We can verify that 32 does not behave similarly to any tones on CVV and CVR.

(20) Xiamen tone sandhi:

a. On CVV and CVR:

$$\begin{tikzcd} 53 \arrow[r] & 44 \arrow[r] & 22 \arrow[dl] & 24 \arrow[l] \\ & 21 \arrow[ul] & & \end{tikzcd}$$

b. On CVO:

4	→ 21	
32	→ 4	for syllables ending in p, t, k
	→ 53	for syllables ending in ʔ

In Changzhou (Wang 1988), a northern Wu dialect of Chinese, five tones can occur on CVV and CVR—55, 13, 523, 24, and 45, and two tones can occur on CVO—23 and 5. The small rise 23 on CVO looks like an incomplete phonetic realization of either 13 or 24, which can occur on CVV and CVR. But if we look at the tone sandhi behavior of Changzhou, shown in (21), we can see that 23 does not behave similarly to either 13 or 24. In the table, tones on CVO are underlined for easy identification.

(21) Changzhou tone sandhi:

$\sigma_1 \backslash \sigma_2$	$\underline{23}$	$\underline{5}$	55	13	45
$\underline{23}$	$\underline{2}$-$\underline{5}$		$\underline{1}$-13		
13	11-$\underline{3}$		11-33		
24	11-$\underline{24}$		11-24		

These examples illustrate that the smaller pitch excursion on CVO cannot always be the result of phonetic implementation. In other words, it cannot be taken as the phonetic reduction of contour tones that can occur on CVV and CVR, since these tones behave independently from other tones in phonological processes such as tone sandhi. Therefore, it is up to the phonology to rule out pronounced pitch excursions on CVO syllables, not just phonetic implementation. Moreover, Zhang (1998) shows that the durational properties of syllables, such as the shortness of CVO, can play a role in determining the sandhi behavior of the tones they carry (see Zhang 1998 for accounts of Yangqu, Shuozhou, and Changzhou tone sandhi). This also indicates that these properties cannot only be left in the realm of phonetics; they are relevant to phonological patterning and thus must be accessible to the phonological grammar.[5]

2.5 Local summary

I have shown in this section that crosslinguistic patterns of contour tone licensing cannot be fully captured by the limited mechanism available to moraic theory. Three arguments are given: Contour tones are sometimes preferentially licensed in positions

[5] Let us note that how the grammar eventually produces different tonal inventories on different syllable types is a separate question, and it is not clear what the answer would be, since under the "Richness of the Base" tenet in Optimality Theory (Prince & Smolensky 1993), an underlying tone that surfaces faithfully in tonally less advantaged syllables (e.g., CVO) should also be a possible UR on tonally advantaged syllables (e.g., CVV, CVR), and the analysis I propose in §3 predicts that it should surface intact as well. The purpose of the discussion here is to point out that this is a problem that must be dealt with in phonology, and low-level phonetic implementation does not help to provide the answer. Constraints on minimum perceptual distance between a contrasting pair *à la* Flemming (1995) may be a fruitful direction in which the answer can be sought.

that are not necessarily moraically privileged; the moras run out of scale when more than three levels of contour licensing ability have to be distinguished; and contour tones with the same number of pitch targets sometimes have different positional restrictions. I have also shown that these licensing properties are all related to the sonorous duration of the rime: prosodic-final syllables, syllables in shorter words, CVV and CVR are longer than *ceteris paribus* non-final syllables, syllables in longer words, and CVO, respectively; consequently, not only are they better licensers for contour tones in general, they are also better licensers for those contour tones that are inherently time-consuming, such as rising tones, tones with pronounced pitch excursion, and tones with more than two pitch targets.

3. A direct approach

3.1 Overview of the theoretic apparatus

In light of the difficulties that moraic licensing faces, I propose an approach to contour tone licensing, which I term the "direct approach", which dispenses with the mora and allows phonetically more informed scales to enter into phonological representations.

One such scale is the Canonical Durational Category (CDC) of a rime. It represents the duration of the sonorous portion of a rime in the canonical speaking rate and style. It is categorical in the sense that two rimes belong to two different CDCs only if there are systematic factors, such as being in prosodic-final position, that influence their duration to a degree safely perceivable by listeners. Crucially, it serves as a phonetic index for a syllable's contour tone bearing ability.[6] The second such scale is what I term the Tonal Complexity of a contour tone, which is defined relationally as follows: for any two tones T_1 and T_2, let D_1 and D_2 be the minimum CDC values required for the production and perception of T_1 and T_2 respectively. T_1 is of higher Tonal Complexity than T_2 if and only if $D_1>D_2$.

The definition of CDC and Tonal Complexity crucially depends on the assumption of canonicality. This is necessary as syllable duration and pitch range vary under different speaking rates and styles, and the "tolerance level" for tone slope varies too. Since the standard mode of speech is what language users are most frequently exposed to and most frequently utilize, it is reasonable to assume that this is the mode that defines these

[6] In Zhang (2002), the phonetic index for contour tone bearing ability of a syllable is $C_{CONTOUR}$, which incorporates the difference in sonority between a vowel and a sonorant consonant. I simplify the discussion here by only referring to the duration, as the sonority difference is not the main concern of this paper.

quantitative values (see Steriade 1999 for a similar assumption). What necessarily comes with the assumption of canonicality is one of normalization, as only under this assumption can we discuss the grammatical behavior of different rates and styles and account for the stability of the phonological system across these rates and styles. This assumption is justified by experimental evidence that speakers can attend to and compensate for fluctuations in speaking rate and style. For instance, many perceptual studies show that the speaking rate of the stimuli influences listeners' perceptual boundary between two segments if this boundary is dependent on duration (Port 1979, Miller & Grosjean 1981, Pols 1986). For studies on tone normalization, see Leather (1983), Moore (1995), and Moore & Jongman (1997).[7]

Now we can define the markedness constraint family $*T_i$-R_j as in (22). It bans different contour tones on different types of rimes. There are two sets of universal rankings for this constraint family, as in (23). Both sets are driven by phonetic rationale—(23a) by the fact that a rime with a shorter CDC is less good a contour carrier than a rime with a longer CDC, (23b) by the fact that a contour with higher Tonal Complexity is less likely to surface on a rime than a contour with lower Tonal Complexity.[8]

(22) $*T_i$-R_j: contour tone T_i cannot occur on rime R_j.

(23) a. If $CDC(R_n)>CDC(R_m)$, then $*T_i\text{-}R_m \gg *T_i\text{-}R_n$.

b. If $TC(T_x)>TC(T_y)$, then $*T_x\text{-}R_j \gg *T_y\text{-}R_j$.

To visualize the effects of Tonal Complexity and CDC on the ranking of these

[7] Alternatively, the effects of speech rate may be dealt with *à la* Flemming (2001), who considers it necessary for a phonological system to specify a set of contrasting sounds that represents an optimal balance between articulatory and perceptual constraints *across a wide range of speech rates*. The design of a phonological grammar is like the design of an engine, which requires fuel-efficiency not only at 55 miles per hour, but across a range of speeds. Therefore, the phonetic realization of each form must not be fixed, but a function of speech rate, and the evaluation of articulatory and perceptual constraints must be defined over these variable realizations. The canonicality assumption is a simpler one from the grammatical perspective, as the grammatical computations involve only one speech rate. But it is a more drastic assumption empirically—it is less than obvious why only one speech rate should be used in determining the contrastive system if the system is actually used at various speech rates (Flemming 2001:31-32, fn.12); and it is in fact not clear that there is really one rate that can be identified as the "canonical rate." There is not enough evidence at this point to opt for either of these two positions.

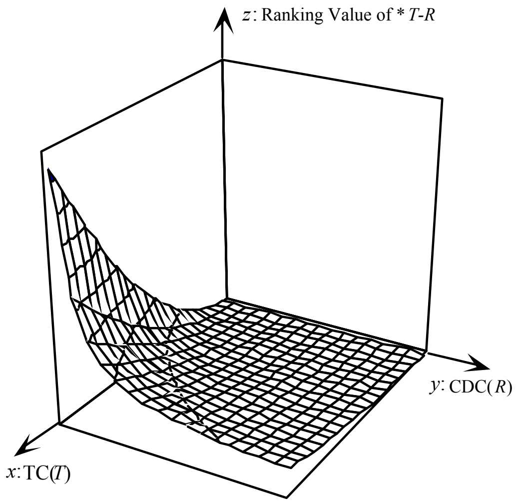
[8] In Zhang (2002), the definition of the tonal markedness constraints directly refers to the Tonal Complexity of the contour and the contour-bearing ability (i.e., $C_{CONTOUR}$, see fn.5) of the syllable. It is my contention that the constraints must refer directly to such phonetic scales to allow more restricted and yet more accurate predictions, and this point has been argued for elsewhere (e.g., Zhang 2002, in press).

constraints, let us assume that every constraint is associated with a *Ranking Value*, with a higher *Ranking Value* indicating a higher constraint ranking. Then the *Ranking Value* of the constraint *T-R can be considered a function of the Tonal Complexity of T—TC(T), and the CDC value of R—CDC(R), as shown in (24).

(24) *Ranking Value* of *T-R = f_{RV}(TC(T), CDC(R))

From (22) and (23), we know that f_{RV} monotonically increases when TC(T) increases, but monotonically decreases when CDC(R) increases. The function f_{RV} can thus be schematically represented in a 3-D space as in (25).[9]

(25) The *Ranking Value* of *T-R as a function of TC(T) and CDC(R):



The graph in (25) is only a schematic. Crucially, for two constraints whose relevant components do not stand in the relationships described in (23), and no ranking between the two constraints can be deduced by transitivity through a third constraint, no claim

[9] One anonymous reviewer raises the question of how a level tone interacts with CDC(R). The position taken here is that a level tone has the lowest Tonal Complexity and therefore projects the lowest ranking value for *T-R. Consequently, the answer to why the 1st tone in Mandarin Chinese (high level) has a longer duration than the 4th tone (falling) must be sought elsewhere. The contour-specific durations predicted by the mechanism here are *necessary* conditions for the realization of these contours. Provided that these conditions are met, languages are free to impose additional durational requirements on a subset of the tones in the inventory. For Mandarin Chinese, the 1st tone has apparently been awarded extra duration, and the speakers of Mandarin need to learn the durational pattern independently from the patterns informed by the phonetic scales.

regarding the intrinsic ranking between them has been made, and their ranking should be determined on a language-specific basis. In other words, $*T_i\text{-}R_i$ and $*T_j\text{-}R_j$ are intrinsically ranked only under the following three conditions: (a) $TC(T_i)=TC(T_j)$; (b) $CDC(R_i)=CDC(R_j)$; (c) $TC(T_i)>TC(T_j)$ and $CDC(R_i)<CDC(R_j)$. The general claim here is that intrinsic rankings can only be determined locally or transitively. This has been proposed as the *Local-Ranking Principle* by Boersma (1998).

The licensing mechanism in (22) and (23) is drastically different from moraic licensing in that it encodes the phonetic properties of the rime crucial to contour licensing and the complexity of contours including the direction and magnitude of pitch change. In so doing, it permits more levels of distinction to be made and a closer tie between rime types and contour tone types, both of which prove to be necessary by the survey of contour tone distribution.

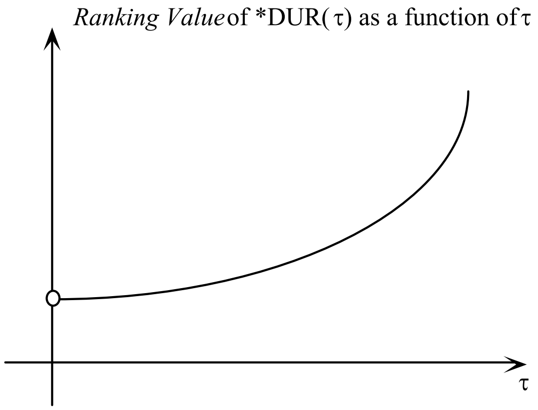
When a contour tone encounters a short sonorous rime, contour flattening and rime lengthening can both alleviate the problem. But these fixes should not come at no cost. To prevent free lengthening of the rime, we first assume that for every segment *x* in a prosodic environment independent of tone, there is a minimum duration associated with it under the canonical speaking rate and style, and these minimum duration requirements must be met. The prosodic environment here includes segment length, stress, proximity to prosodic boundaries, number of syllables in the word, etc. We can then define constraint *DUR(τ) as in (26a), and posit that all *DUR constraints are universally ranked as in (26b).

(26) a. *DUR(τ): for a rime in a prosodic environment, the cumulative duration in excess of the minimum duration for each segment cannot be τ or more. (τ>0)

b. If $\tau_i>\tau_j$, then $*\text{DUR}(\tau_i) \gg *\text{DUR}(\tau_j)$.

If we consider the ranking value of *DUR(τ) to be a function of τ (τ>0), with a higher *Ranking Value* indicating a higher ranking, then according to (26), this function is monotonically increasing, as shown schematically in (27).

(27)



The underlying tonal shape of the contour is protected by a family of tonal faithfulness constraints defined on a discrete scale from *1* to a hypothetical value *n* as in (28). The insight here is to acknowledge that a farther deviation from the input tone is a worse violation of faithfulness, and such violation is penalized by a higher ranked faithfulness constraint. The distance from the input is measured by a perceptual scale, which I do not spell out due to space limitations. But see Zhang (2002) for specific discussion on tones, and Steriade (2001) for general discussion on universal rankings of faithfulness constraints projected from the "Perceptual Map".

(28) a. $\forall i$, $1 \leq i \leq n$, $\exists$ constraint PRES(T, *i*), defined as:
An input tone T_I must have an output correspondent T_O, and T_O must be less than *i* steps away on the perceptual scale from T_I.
b. PRES(T, *n*) » PRES(T, *n-1*) » ... » PRES(T, *2*) » PRES(T, *1*)

3.2 Sample analysis

With the formal apparatus in place, let us consider its predictions. Suppose that in language *L*, there exists an underlying contour tone T with a pitch excursion of Δf under the standard speaking rate and style. Let us see what predictions the apparatus makes when the contour encounters a rime R whose CDC value is *c* and whose minimum sonorous rime duration is *d*. The predicted input-output mapping may be either alternation or allophonic distribution.

3.2.1 No change necessary

The first possibility is that the PRES(T) and *DUR constraint families outrank *T-R *en masse*. Under this ranking, the contour faithfully surfaces on the given rime without lengthening. This is because any flattening of the contour or lengthening of the sonorous rime duration in order to satisfy *T-R will incur violations in the higher ranking PRES(T) or *DUR constraint families, as illustrated by the tableau in (29).

(29) $T_{\Delta f}, R_d \rightarrow \Delta f, d$

	$T_{\Delta f}, R_d$	PRES(T)	*DUR	*T-R
☞	faithful: $\Delta f, d$			*
	contour reduction: $\Delta f\text{-}f_0, d$	*!		
	rime lengthening: $\Delta f, d+d_0$		*!	

This ranking also predicts that on a rime R' with a CDC value greater than *c*, Δf will also be faithfully realized, since the constraint *T-R' will be even lower ranked than *T-R. This pattern is attested in !Xũ (Snyman 1970), ǂKhomani (Doke 1937), and a number of Chinantec languages which allow all tones on any syllable, be it open or checked, long-voweled or short-voweled. Although most of the sources I consulted on these languages do not give phonetic details of tone and duration, thus it is possible that the contour tones on shorter syllable types are somewhat flattened, or these syllables are somewhat lengthened, there is some phonetic documentation on Lalana Chinantec (Mugele 1982) showing that the same contour tone exhibits relative stability of onset and endpoint on different syllable types, and the same syllable type exhibits relatively stable duration when carrying different tones.

3.2.2 Partial contour reduction

The second possibility is that *T-R outranks some, but not all PRES(T) constraints, but the *DUR constraint family *en masse* is still undominated. Under this ranking, the contour is flattened to satisfy the *T-R constraint, but no extra duration can be added to the sonorous portion of the rime, as illustrated in (30).

(30) $T_{\Delta f}, R_d \rightarrow \Delta f\text{-}f_0, d$

$T_{\Delta f}, R_d$	*DUR	*T-R	PRES(T)
faithful: $\Delta f, d$		*!	
☞ contour reduction: $\Delta f\text{-}f_0, d$			*
rime lengthening: $\Delta f, d+d_0$	*!		

This ranking also predicts that on a rime R' with a CDC value greater than c, Δf will be more faithfully realized, i.e., realized with less or no reduction of the pitch excursion. This is because the relevant constraint *T-R' will be lower ranked than *T-R, which will allow more PRES(T) constraints to exert influence on the output form. This characterizes the pattern in which certain contour tones can have a full realization on syllables with greater CDCs, but are partially flattened elsewhere. Pingyao Chinese's flattening of 53 and 13 on CVO syllables to 54 and 23 is an example of this sort.[10]

3.2.3 Complete contour reduction

The third possibility is to have all *T_x-R and *DUR constraints outrank all the relevant PRES(T) constraints. That is, *δ-R, where δ represents the smallest perceptible pitch excursion, outranks the PRES(T, *i*) constraint that penalizes changing the tone T to a level tone. This ranking predicts that the tone T will be flattened all the way to a level tone, as illustrated in (31).

[10] Pingyao's pronounced contours (53 and 13) behave identically when they dock onto CVO syllables in that they both reduce partially. It is also possible for the system to predict that only one of them reduce: if we take into account the fact that rising is more time-consuming than falling, then we may predict that 13 needs to reduce, but 53 does not; if the underlying falling contour is of greater excursion than the rising, then the falling may need to reduce while the rising may not.

(31) $T_{\Delta f}, R_d \rightarrow 0, d$

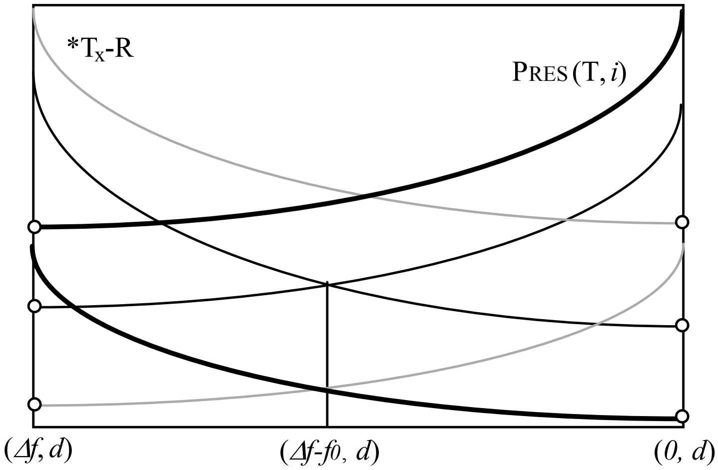
	$T_{\Delta f}, R_d$	*DUR	*δ-R	PRES(T, *i*)
	faithful: Δf, d		*!	
	partial contour reduction: $\Delta f\text{-}f_0$, d		*!	
☞	complete contour reduction: 0, d			*
	rime lengthening: Δf, $d+d_0$	*!		

This ranking still predicts that on a rime R' with a CDC value greater than *c*, Δf will be more faithfully realized: *δ-R' will be lower ranked than *δ-R, which will allow more PRES(T) constraints to exert influence on the output form. This represents the most commonly attested pattern of contour tone restrictions in languages, i.e., certain contour tones cannot occur on syllables with low CDC values. We have seen many examples of this sort, e.g., Xhosa's restriction of contour tones to stressed syllables, Beijing Chinese' restriction of 213 to word-final syllables, Mende's restriction of L͡HL to monosyllabic words with a long vowel, etc.

3.2.4 Interim summary

The scenarios described in §3.2.1—§3.2.3 can be summarized in the schematic graph in (32). In the graph, the *x*-axis represents tonal candidates. Since all *DUR constraints are always ranked on the top tier in the scenarios described so far, I only consider candidates that respect these constraints, i.e., candidates with no lengthening. The leftmost candidate on the *x*-axis is the most faithfulness to the input, with no flattening at all—(Δf, d). The rightmost candidate is the one with complete flattening—(0, d). d is the sonorous rime duration of the candidate rime, and it is the same in all the candidates considered here. The *y*-axis represents constraint ranking—the higher the *y* value, the higher the ranking. The curves in the graph represent the highest ranked constraints in the *T_x-R and PRES(T, *i*) families that the candidates on the *x*-axis violate.

(32) Interaction between $*T_x$-R and PRES(T, *i*) yielding different degrees of contour reduction:



The thick black lines in the graph indicate the ranking of the two constraint families that produces the faithful realization of the pitch excursion Δf, which is the leftmost candidate on the *x*-axis. The highest ranked constraint it violates is *T-R. Any other candidate to the right deviates from the input, and hence will violate a higher ranked PRES(T, *i*) constraint.

The thin black lines indicate the ranking that produces partial reduction of the contour to $\Delta f\text{-}f_0$, which is the candidate on the *x*-axis that corresponds to the point of intersection of the two curves. Any candidate to the left violates a higher ranked $*T_x$-R constraint, and any candidate to the right violates a higher ranked PRES(T, *i*) constraint.

The gray lines indicate the ranking that forces complete reduction of the contour tone to a level tone, which is the rightmost candidate on the *x*-axis. The highest ranked constraint it violates is the highest ranked PRES(Tone, *i*) constraint. Any other candidate to the left reduces the contour less, and will violate a higher ranked $*T_x$-R constraint.

3.2.5 Non-neutralizing lengthening

The fourth possibility is that *T-R outranks some *DUR constraints, but the PRES(T) constraint family *en masse* is undominated. Under this ranking, the tone-bearing portion of the rime is lengthened to satisfy the *T-R constraint, but the contour must be faithfully realized, as illustrated by the tableau in (33).

(33) $T_{\Delta f}, R_d \rightarrow \Delta f, d{+}d_0$

$T_{\Delta f}, R_d$	PRES(T)	*T-R	*DUR
faithful: $\Delta f, d$		*!	
contour reduction: $\Delta f\text{-}f_0, d$	*!		
☞ rime lengthening: $\Delta f, d{+}d_0$			*

This ranking also predicts that on a rime R' with a greater CDC value, there will be a lesser degree of lengthening or no lengthening at all depending on what the sonorous rime duration is. This pattern does not seem prevalent in the survey, but this may be due to the fact that the primary attention has been devoted to documenting the restrictions of contour tones on certain syllable types in the data sources, so when a syllable type is able to carry a certain contour, the durational change of the syllable is considered a phonetic side-effect and has escaped the attention of many. We do have a few examples in which this pattern is instantiated. E.g., in Mitla Zapotec (Briggs 1961), a rising tone lengthens the duration of its carrier, and in Wuyi Chinese (Fu 1984), a CVO syllable is drastically lengthened to carry a complex contour 213.

3.2.6 Neutralizing lengthening

It is also possible that the lengthening is neutralizing. Let us suppose that the minimum durations for a short vowel and a long vowel are *d* and *2d* respectively. Then when *T-$R_{2d-\delta}$ (with δ being a very short duration) outranks *DUR(d), while all PRES(T) constraints are still ranked on top, the ranking predicts neutralizing lengthening when the tone T occurs on a short vowel. This is illustrated in the tableau in (34). The first candidate, with no contour flattening and no lengthening, violates the highly ranked *T-$R_{2d-\delta}$; the second candidate, with contour flattening, violates at least one of the highly ranked PRES(T) constraints. The third candidate, with insufficient lengthening, still violates the constraint *T-$R_{2d-\delta}$. The last candidate, with sufficient lengthening, only violates the lowly ranked *DUR constraints, and is therefore the winner.

(34) $T_{\Delta f}, V_d \rightarrow \Delta f, V_{2d}$

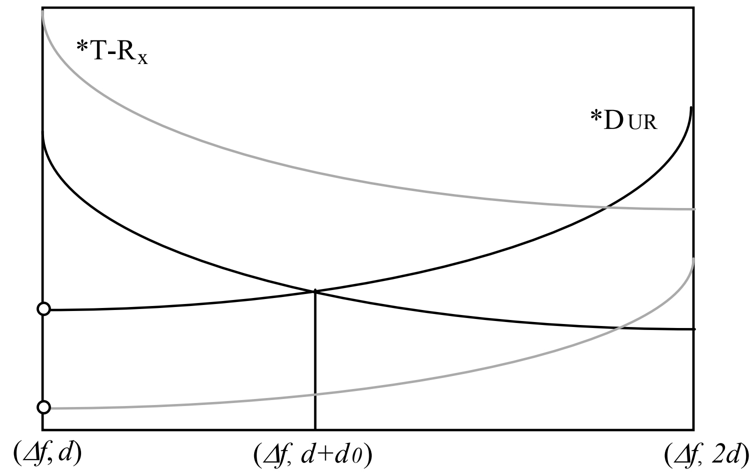
$T_{\Delta f}, V_d$	PRES(T)	*T-R$_{2d\text{-}\delta}$	*DUR(d)
$\Delta f, V_d$		*!	
$\Delta f\text{-}f_0, V_d$	*!		
$\Delta f, V_{2d\text{-}\delta}$		*!	
☞ $\Delta f, V_{2d}$			*

This ranking also predicts that on a long vowel, the tone T can be faithfully realized. This pattern is attested in Gã. There is a vowel length contrast in this language. But when a rising tone co-occurs with a short vowel due to morphological concatenation, neutralizing lengthening (Paster 1999).

3.2.7 Interim summary

The scenarios described in §3.2.5—§3.2.6 can be summarized in the schematic graph in (35). In the graph, the *x*-axis represents durational candidates. Since all *PRES(T) constraints are always ranked on the top tier in these scenarios, I only consider candidates that respect these constraints, i.e., candidates with no contour reduction. The leftmost candidate on the *x*-axis is the most faithful to the input, with no lengthening at all—(Δf, d). The rightmost candidate is the one with neutralizing lengthening—(Δf, $2d$). The *y*-axis represents constraint ranking—the higher the *y* value, the higher the ranking. The curves in the graph represent the highest ranked constraints in the *T-R$_x$ and *DUR families that the candidates violate.

(35) Interaction of *T-R$_x$ and *DUR yielding different degrees of lengthening:



The black lines in the graph indicate the ranking of the two constraint families that produces partial lengthening of the vowel to $d+d_0$, which is the candidate on the *x*-axis that corresponds to the point of intersection of the two curves. Any candidate to the left violates a higher ranked *T-R_x constraint, and any candidate to the right violates a higher ranked *DUR constraint.

The gray lines indicate the ranking that forces neutralizing lengthening, which is the rightmost candidate on the *x*-axis. The highest ranked constraint it violates is the highest ranked *DUR constraint. Any other candidate to the left lengthens less from the input, and will induce the violation of a higher ranked *T-R_x constraint.

3.2.8 Contour reduction + rime lengthening

The last possibility is that *T-R outranks some *DUR constraints and some PRES(T) constraints. Under this ranking, the avoidance of the *T-R constraint violation is achieved by contour reduction and rime lengthening simultaneously.

To illustrate this, let us assume the following: $f_0>f_1$, $d_0>d_1$, and $\Delta f\text{-}f_0$ and $\Delta f\text{-}f_1$ are *i* and *j* steps away on the perceptual scale from tone T (which has the pitch excursion Δf) respectively. Given that $f_0>f_1$, we know that $i>j$, meaning that $\Delta f\text{-}f_1$ is perceptually closer to tone T than $\Delta f\text{-}f_0$. Based on the intrinsic rankings among the *DUR and PRES(T) constraint families, these relations render the intrinsic rankings shown in (36).

(36) a. *DUR(d_0) » *DUR(d_1)
  b. PRES(T, *i*) » PRES(T, *j*)

If the relevant *T-R is ranked on a par with *DUR(d_0) and PRES(T, *i*), but outranks *DUR(d_1) and PRES(T, *j*), then the winning candidate will have a flattened contour $\Delta f\text{-}f_1$ and a lengthened duration $d+d_1$. Just flattening the contour to satisfy *T-R is too costly for the PRES(T) constraint family as it incurs a violation of the highly ranked PRES(T, *i*); and just lengthening the rime is too costly for the *DUR constraint family as it incurs a violation of the highly ranked *DUR(d_0). The tableau in (37) illustrates these arguments.

(37) $T_{\Delta f}, R_d \rightarrow \Delta f\text{-}f_1, d+d_1$

$T_{\Delta f}, R_d$	PRES(T, *i*)	*DUR(d_0)	*T-R	PRES(T, *j*)	*DUR(d_1)
faithful: $\Delta f, d$			*!		
lots of contour reduction: $\Delta f\text{-}f_0, d$	*!			*	
lots of rime lengthening: $\Delta f, d+d_0$		*!			*
☞ some reduction, some lengthening: $\Delta f\text{-}f_1, d+d_1$				*	*

This ranking also predicts that on a rime R' with a duration longer than *d*, there will be a lesser degree of flattening, or a lesser degree of lengthening, or both, depending on the ranking among the lower-ranked $*T_x\text{-}R_y$, PRES(T), and *DUR constraints. This pattern is instantiated by Hausa, which shows both partial contour flattening and rime lengthening when a CVO syllable carries a falling contour. The factorial typology clearly predicts many variations of this pattern, but this pattern does not seem prevalent in the survey. An explanation is surely needed. I again conjecture that this might be due to the close-to-exclusive attention to the distributional facts about contours and the lack of detailed phonetic documentation of many languages. Upon closer scrutiny of the phonetic realization of tonal contours and duration of rimes that carry them, many such patterns might emerge and the range of variation predicted by the typology can be tested against these phonetic data.

3.3 Summary

In this section, I have proposed an explicit formalization of the direct approach to contour tone distribution and discussed the patterns that are predicted by the model. The factorial typology of the constraints proposed yields six possible outputs for contour tone T on rime R, as summarized in (38). All these patterns are attested.

Couched within Optimality Theory, the approach taken here allows phonetically more informed representations to enter into phonological constraints, and the generalizations regarding contour tone bearing fall out from phonetically motivated constraint rankings. The Tone-Bearing Unit in this approach is the sonorous portion of the rime rather than the mora—a position closer to Yip's. The contour tone licensing ability of a syllable is not due to the mora count, but the Canonical Durational Category it belongs to, and the fact that different contour tones may have different licensing environments is not due to the different mora counts they require, but their Tonal Complexity.

Moreover, as we have seen, the incorporation of normalized phonetic detail in phonology still allows us to predict six distinct patterns of interaction between contour tone and sonorous rime duration, including ones with neutralizing lengthening or shortening. These patterns are produced by the high-ranking of markedness constraint and the low-ranking of either the markedness constraint against duration or tonal faithfulness constraint. An additional mechanism to produce neutralization is to incorporate constraints on the perceptual distance of contrast into the system. This position is explicated in Flemming (2001) and endorsed by Steriade (2001) and Zhang (2000b). With contrast constraints, the markedness constraint hierarchy does not need to be pushed all the way high in the hierarchy to predict neutralization; provided that they force enough lengthening or flattening so that the cost of violation of contrast constraints outweighs the benefit of keeping the contrast, neutralization will happen. And as I show later in §4.3, contrast constraints are independently necessary to capture the reduction of tonal inventory on shorter duration.

Given the phonetically-informed nature of the constraints, the complete factorial typology of the system will show that there is a large number of variations for some of the patterns predicted. This rightfully raises concerns on the explanatory power of the theory. Specifically, the question is to what extend the system over-generates. As Zhang (2000b) has argued, with more detailed phonetic studies, we may realize that many patterns that seem to be over-generated by the factorial typology are in fact attested. A growing body of phonetic literature has shown that many phonetic processes that were thought to be universal exhibit cross-linguistic variations, and these variations usually tie into the system of contrast in the language in question (Keating 1988a, b, Keating & Cohn 1988, Manual 1990, Flemming 2001). It is then possible that there is a better match between the predictions of the factorial typology and the attested patterns than one originally thought.

(38) Outputs of $T_{\Delta f}$, R_d generated by the factorial typology:

Output	Constraint ranking	Example languages
a. Faithful: Δf, d	PRES(T), *DUR ⇓ *T-R	Lalana Chinantec, !Xũ, ǂKhomani
b. Partial contour reduction: Δf-f_0, d	*DUR, *T-R ⇓ some PRES(T)	Pingyao Chinese
c. Complete contour reduction: 0, d	*DUR, *δ-R ⇓ PRES(T, i)	Xhosa, Beijing Chinese
d. Non-neutralizing lengthening: Δf, $d+d_0$	PRES(T), *T-R ⇓ some *DUR	Mitla Zapotec, Wuyi Chinese
e. Neutralizing lengthening: Δf, $2d$	PRES(T), *T-$R_{2d-\delta}$ ⇓ *DUR(d)	Gã
f. Reduction and lengthening: Δf-f_1, $d+d_1$	some *DUR, some PRES(T), *T-R ⇓ some other *DUR, some other PRES(T)	Hausa

4. Further theoretical implications

As further implications of the direct approach, I discuss in this section how it can potentially resolve two conundrums in moraic theory—moraic inconsistency and what I term “non-moraic consistency”, and how it opens the door for the understanding of tonal inventory restrictions on different syllable types.

4.1 Moraic inconsistency

As mentioned earlier, moraic theory generally assumes that in a language, all weight-related phenomena are accounted for by the same moraic representation. Although this finds support in languages like Cairene Arabic, in which stress, word-minimality, and vowel shortening converge on the same moraic representation (McCarthy & Prince 1986), it is found to be problematic in languages like Lithuanian, Classical Greek, Tübatulabal, Yawelmani, and other languages by Hyman (1985), Archangeli (1991), Crowhurst (1991), Steriade (1991), Broselow (1995), and most recently, Gordon (1999).

In Lithuanian (Steriade 1991), monosyllabic roots consisting of only a short open syllable are not allowed; but syllables closed by an obstruent coda, such as *lip* 'rise, climb' are sufficient to satisfy the root minimality requirement. This indicates that if the minimality requirement is two moras, then an obstruent coda must be counted as moraic. But Zec (1988) argues that if we look at other weight-related processes in the language, an obstruent coda should not be counted as moraic. First, in accent distribution, the rising tone accent can only occur on CVV and CVR syllables, but not on CVO. Second, in the formation of infinitive verbs, there is a requirement for the stem to be bimoraic. The vowel is lengthened in CVO stems, but it remains short in CVR stems, indicating that CVR stems are bimoraic, while CVO stems are not. Third, a long vowel is shortened when it is followed by a tautosyllabic sonorant, but not when it is followed by a tautosyllabic obstruent.

In Classical Greek (Attic) (Steriade 1991), CVCC is as heavy as CVVC and CVV for recessive accent assignment, quantitative meter, and word minimality requirement, indicating that the final consonants in CVCC must contribute at least one mora to the syllable. But from the distribution of a High tone that appears on the last syllable of words followed by enclitics, Steriade argues that only vowels are tone-bearing segments in Classical Greek. Her argument goes as follows: the placement of the High tone is blocked when the word has penultimate accent and the penult is either CV or CVC, and this is due to the OCP, which disallows two adjacent High tones; but when the word has penultimate accent and the penult is CVV, the High tone surfaces on the final syllable, and this is because the second vowel in the penult carries a Low tone, which breaks up the High-High sequence. The examples in (39) show that the High tone surfaces when the penult is CVV, but it is blocked when the penult is CV or CVC.

(39) a. High tone surfaces:

óikos	'house'
óikós tis	'some house'
dóoron	'gift'
dóorón tis	'some gift'

b. High tone blocked:

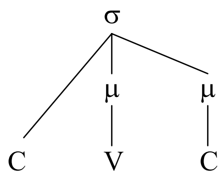
phílos	'friend'
phílos tis	'some friend'
éntha	'there'
éntha te	'and there'

In Yawelmani, Archangeli (1991) shows that mapping a CVC root to a bimoraic morphological template results in the lengthening of the vowel, which indicates that the coda consonant is non-moraic; but long vowels shorten in closed syllables, which could be interpreted as a bimoraic limit on the syllable and consequently leads to the conclusion that the coda consonant is moraic.

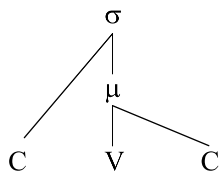
Various proposals have been made to deal with the moraic inconsistency problem, most notably, rule ordering and multileveled representations.

For example, for Classical Greek, Hyman (1985) proposes a margin creation rule, which applies after the accent assignment but before the mapping of the High tone, changing the representation in (40a) to (40b), i.e., associating the coda consonant to the mora contributed by the vowel and removing its own mora. This rule ordering ensures that the coda consonant is moraic in accent assignment, but non-moraic in High tone mapping.

(40) a. Before the margin creation rule:



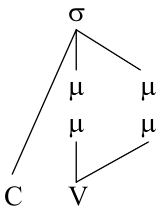
b. After the margin creation rule:



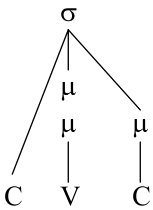
Archangeli (1991) proposes a similar solution to the moraic inconsistency in Yawelmani. She orders Weight-by-Position, which assigns a mora to a coda consonant, after templatic mapping, but before vowel shortening, thus accounting for both the lengthening and the shortening.

Hayes (1995), on the other hand, proposes that moras form a grid within the syllable, with the height of the column determined by the sonority of the segment it is associated with. A sample set of moraic representations for CVV, CVC, and CV in a language that involves moraic inconsistencies of the coda consonant is given in (41). In this conception, processes that treat CVC as bimoraic refer to the lower layer of the grid, while processes that treat CVC as monomoraic refer to the higher layer of the grid.

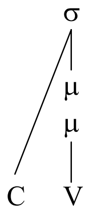
(41) a. CVV



b. CVC



c. CV



Adopting a claim in Steriade (1991), Hayes further conjectures that syllable-external prosodic requirements such as footing, word minimality, and tonal docking generally refer to the higher layer, while syllable-internal requirements such as mora population limits generally refer to the lower layer.

My major objection to the rule-ordering approach is its arbitrariness. Given that there is no *a priori* principle that states which rules should apply before which other rules, it is equally likely for the margin creation rule, e.g., to occur before stress assignment but after tone mapping, and before tone mapping but after stress assignment. Therefore the theory does not predict any asymmetry among processes in treating the weight of a syllable type. It is just as likely for CVO to be considered heavy for tone but light for stress as the opposite. But this turns out not to be true. Gordon (1999) has pointed out that it is much more likely for CVO to be counted as heavy for stress than for tone. His survey shows that, of 41 languages with weight-sensitive contour tone distribution, only two of them (4.8%) treat CVO as heavy; all others require either CVV or CVR for contour tones to surface. But of 69 languages with weight-sensitive stress, 28 of them (40.6%) treat CVO as heavy—a much higher percentage than weight-sensitive tone. Gordon's result on contour tones is corroborated by the survey in Zhang (2002): in 187 languages included in the survey, a total of 104 languages require either CVV or CVR for contour tones to be realized, while only four languages allow contour tones on CVO.

Hayes' solution to the problem does make predictions about the correlation between processes and the level of moraic projection they refer to. But given that stress assignment and contour tone distribution should both be considered syllable-external processes, we are still left without an explanation for the asymmetry between these two processes in their treatment of the CVO syllables.

I believe that the different treatment of CVC, especially CVO, among different weight-related processes lies in the different phonetic requirements of these processes. This line has been explicitly pursued by Gordon (1999). He lays out the possible phonetic bases for six weight-related processes—quantitative stress assignment, contour tone licensing, compensatory lengthening, metrics, syllable templates, and word minimality, as summarized in (42), and argues that these phonetic bases are the driving forces for the phonological patterning of these processes. In particular, he argues that it is the total energy of the rime that determines the ability of the syllable to attract stress, but it is the total *sonorant* energy of the rime that is crucial for its contour tone bearing ability. The fact that it is more frequent for the world's languages to treat CVO as heavy for stress than for tone is determined by the necessity of sonorancy (i.e., presence of energy in the second to fourth harmonics) for tonal perception, but not for stress.

(42) Different weight-related processes and their phonetic considerations:

Weight-related processes	Phonetic bases
Quantitative stress assignment	Total perceptual energy of the rime[11]
Contour tone licensing	Total sonorant energy of the rime
Compensatory lengthening	Rime duration
Metrics	Rime duration[12]
Syllable templates	Syllable isochrony
Word minimality	CVV, CVC ok: duration CVV ok: support of a minimal intonational contour[13]

Without carrying the discussion too far afield, let us simply note that the intuition provided by Gordon can be readily formalized in the direct approach, as the approach does not rely on a unitary moraic representation, but incorporates process-specific phonetic representations.[14]

---

[11] The total perceptual energy of the rime is calculated by Gordon as follows. First, the average amplitude (RMS) in decibels of the target vowel and coda consonant was calculated relative to a reference vowel. Second, the relative RMS of each segment was converted to a value representing perceived loudness. Third, the relative loudness value for each segment was multiplied by the segment duration, yielding the perceptual energy value of the segment. Finally, the perceptual energy values of the rime segments are added together, yielding the total perceptual energy of the rime (Gordon 1999:170).

[12] Gordon (1999) does not specifically discuss the phonetic basis for the heavy/light distinction in metrics. But given that the weight criteria for metrics and compensatory lengthening are always consistent with each other within a language (Gordon 1999:248), and that the phonetic basis for compensatory lengthening is rime duration, I assume that the weight criterion for metrics is also dependent on rime duration.

[13] This is only one of the possible phonetic bases that Gordon (1999) provides for word minimality. Other possibilities include: content words must possess sufficient amount of energy to increase their perceptual salience; the total material in a morpheme should be maximized to increase its chances of being recovered from the signal; a short open syllable is disallowed to avoid neutralization in the face of stress, final lengthening, and the greater duration induced by being in a word with fewer syllables.

[14] As discussed in fn.7, in the unabridged version of the direct approach (Zhang 2002, in press), the tonal markedness constraints directly refer to phonetic properties of the syllable ($C_{CONTOUR}$) and the tone (Tonal Complexity). This is what distinguishes my approach from Gordon's. The formal system that Gordon (1999) proposes does not directly encode phonetic detail. Rather, the phonetics is mediated through phonological entities such as the X slot. Therefore, in his account of contour tone distribution, he uses constraints such as the ones in (1), and posits the university constraint ranking (1a) » (1b) » (1c). To account for patterns in Hausa (contours on CVV, CVR,

4.2 Non-moraic consistency

As a mirror-image of moraic inconsistency, non-moraic consistency refers to phonological phenomena that are not weight-related, but have similar behavior to contour tone licensing. One such phenomenon is the licensing of diphthongs, as discussed in Zhang (2001). Phonetically, diphthongs are similar to contour tones in that they also require articulatory movements, and their perception also requires perceiving gradual acoustic changes. Both of these properties determine that diphthongs, just like contour tones, need sufficient duration to be realized. And indeed, the distributional properties of diphthongs are very similar to those of contour tones, in that they are also more likely to be licensed in positions with inherently longer duration, such as stressed or prosodic-final syllables, as shown in Zhang (2001). But this parallel is completely lost if contour tones are licensed by moras projected from the segmental tier, since diphthongs are <u>on</u> the segmental tier and can project moras themselves, so their licensing restrictions cannot be due to the lack of moras. Then no matter where they stem from, their account is necessarily different from that for contour tone restrictions. Hence the

---

CVO), Kiowa (contours on CVV, CVR), and Xhosa (contours on CVV), he only needs to rank FAITH(tone) with respect to the universal hierarchy. E.g., the ranking (1a) » FAITH(tone) » (1b) » (1c) derives the Hausa pattern.

(1)	a.	*T T				A contour tone is licensed
		\ /	unless	$[XX]_R$	:	by a rime containing two
		R				timing slots.
	b.	*T T		$[XX]_R$		A contour tone is licensed
		\ /	unless	\ /	:	by a rime containing two
		R		[+sonorant]		timing slots that are [+sonorant].
	c.	*T T		$[XX]_R$		A contour tone is licensed
		\ /	unless	\ /	:	by a rime containing two
		R		[+syllabic]		timing slots that are [+syllabic].

Three arguments can be made to support the direct encoding of phonetic properties in phonology, contra Gordon. First, as I have shown earlier in this paper, the effects of phonetic duration and sonority on contour tone restrictions far exceeds what Gordon's phonological system in (1) inherently predicts. Second, Gordon's approach in fact does not predict the partial contour flattening and rime lengthening in CVO in Hausa—it predicts that contour tones will surface faithfully on CVO. Thus such partial flattening and lengthening must be taken as purely phonetic by Gordon. And as we have seen in §2.4.3, this position is untenable. Third, the process-specific nature of weight phenomena finds a more direct account if the process-specific phonetic properties (e.g., sonorous rime duration for tone, overall rime energy for stress) are directly encoded in the phonological grammar (e.g., in the form of markedness constraints).

similarities between diphthong and contour tone restrictions are left without an explanation. Under the direct approach however, this parallel is expected, since the need for duration for both of these phonological entities is encoded in the analysis.

4.3 The size of tonal inventory on different syllable types

The final theoretical implication of the direct approach is on its potential to account for the differences in the size of the tonal inventory on different syllable types. If we look back at the Pingyao data in (15), we shall notice that not only do CVʔ syllables have contour tones with smaller pitch excursion, they also have fewer contour tones. The rising contour 35, which can occur on CVV and CVR, has no counterpart in the tonal inventory of CVO. This is a very common phenomenon in Chinese dialects. In those dialects with CVO syllables (which include most of Wu, Min, Jin, Yue, and Hakka dialects), there are usually a maximum of two contrastive tones on CVO, but four to six on CVV and CVR. Often times, the tones that occur on CVO are contour tones, as in Pingyao and Xiamen. So the difference in the size of tonal inventory of different syllable types cannot simply result from a contour vs. level distinction. Then what is the basis for this difference?

The moraic approach does not have much to say about this difference. As long as the structural requirement for a contour tone—two moras—is met on CVO, as it has to be, given the presence of contour tones on this syllable type, the theory itself provides no explanation as to why one contour tone can occur while another cannot.

This is a problem for the direct approach as well. The situation is the same: if the CDC value of a syllable is high enough for one contour tone to surface, why does another contour tone with the same tonal complexity fail to surface? But the direct approach is a phonetically more articulate theory. It allows the phonology to access phonetic detail. One type of phonetic detail that the phonology could conceivably have access to is the perceptual distance between two contrasting phonological entities, and here, the relevant phonological entities are tones. Flemming (1995) and Kirchner (1997) have both proposed to introduce constraints that require a minimum distance between phonological contrasts into the phonological system, Flemming by MINDIST, Kirchner by POLAR. Take MINDIST for instance, it is a series of intrinsically ranked constraints MINDIST=M (MINDIST=1 » MINDIST=2 » MINDIST=3…), which requires phonological contrasts to be *M* “steps” apart. When it is interleaved with another series of intrinsically ranked constraints MAINTAIN-N-CONTRASTS (MAINTAIN-1-CONTRAST » MAINTAIN-2-CONTRASTS » MAINTAIN-3-CONTRASTS…), which requires the maintenance of *N* contrasts, the constraint hierarchy ensures that the resulting members of an inventory are kept a maximum perceptual distance apart from each other. Adopting the MINDIST and MAINTAIN-N-CONTRASTS into the direct approach, we may assume that given the shorter duration on CVO than CVV and CVR, the perceptual distance between the same tones on CVO is smaller than that on CVV and CVR. This determines that we shall only be able to maintain fewer tonal contrasts on CVO than CVV and CVR.

Let us assume that on the canonical duration of CVV or CVR, adjacent tones in 13, 35, and 53 are at a distance of two steps along a linear perceptual scale: 13 and 35 are two steps from each other, so are 35 and 53; 13 and 53 are four steps apart. Intuitively, this is because 13 and 35 differ in average pitch height, 35 and 53 differ in pitch change direction, and 13 and 53 differ in both parameters. On the canonical duration of CVO however, the adjacent tones in 13, 35, and 53 are only at a distance of one step, due to the shortness of the CVO duration. The constraint ranking in (43) will then ensure that 13, 35, and 53 will be the tonal inventory on CVV and CVR, while 13 and 53 will be the tonal inventory on CVO.

(43) MAINTAIN-1-CONTRAST » MINDIST=1 » MINDIST=2 »
MAINTAIN-2-CONTRASTS » MINDIST=3 » MAINTAIN-3-CONTRASTS

The tableaux in (44) show how the inventories are derived. In (44a), since the tones 13-35-53 are two steps apart on the perceptual scale, they only violate the lowest ranked MINDIST constraint here: MINDIST=3; and keeping all of them will only violate the lowest ranked MAINTAIN-N-CONTRASTS constraint here: MAINTAIN-3-CONTRASTS. Having one more tone in the inventory will violate MINDIST=2, and having one fewer tone in the inventory will violate MAINTAIN-2-CONTRASTS, both of which outrank MINDIST=3 and MAINTAIN-3-CONTRASTS. Thus 13-35-53 is the optimal tonal inventory of CVV and CVR. In (44b) however, since 13-35-53 are only one step apart on the perceptual scale due to the short duration, having all of them in the inventory will violate MINDIST=2. Removing 35 from the inventory will result in a violation of MAINTAIN-2-CONTRASTS, but satisfy MINDIST=2. Given that MINDIST=2 » MAINTAIN-2-CONTRASTS, we conclude that 13-53 is the optimal tonal inventory of CVO. Notice that this system is essentially Pingyao's system.

(44) a. On CVV and CVR: 13-35-53

	MAINTAIN 1 CONTRAST	MINDIST =1	MINDIST =2	MAINTAIN 2 CONTRASTS	MINDIST =3	MAINTAIN 3 CONTRASTS
13-53				*!		*
13-35				*!	*	*
☞ 13-35-53					*	*
13-35-55-53			*!		*	

b. On CVO: 13-53

	MAINTAIN 1 CONTRAST	MINDIST =1	MINDIST =2	MAINTAIN 2 CONTRASTS	MINDIST =3	MAINTAIN 3 CONTRASTS
☞ 13-53				*	*	*
13-35			*!	*	*	*
13-35-53			*!		*	*
13-35-55-53		*!	*		*	

As Flemming (1995) has pointed out, the use of MAINTAIN-N-CONTRAST constraints prevents the use of traditional faithfulness constraints, as the grammar should have one and only one mechanism of deriving contrasts. Therefore, the adaptation of this approach to the current system is not straightforward. But the point here is that, given the phonetically rich nature of the current system, it is *possible* for it to adopt such a proposal, which requires reference to phonetic detail, to account for the difference in tonal inventory size of different syllable types. For a purely representational approach based on the mora, it is not clear how this issue can be addressed.

5. Conclusion

In conclusion, I have argued that bimoraicity is neither sufficient nor necessary for contour tone licensing; contour tones are properties of the syllable and the sonorous portion of the rime is the Tone-Bearing Unit. A direct approach, which incorporates the phonetic basis of contour tone realization, captures the crosslinguistic patterns of contour tone distribution and provides a possible explanation for two conundrums in moraic theory—moraic inconsistency and non-moraic consistency and opens the door for the understanding of tonal inventory restrictions on different syllable types.

Of course, the system sketched out above only addresses one aspect of tonal phonology—the relation between syllable duration and the complexity of contour tones. Other factors, such as a glottal coda and the voicing of an onset obstruent are also known to affect tonal realization and need to be incorporated into the complete system of tonal phonology. Furthermore, many complicated tonal alternations, like the tone sandhi systems of Chinese, are admittedly rich in diachronic vestiges that have lost the phonetic underpinnings and must be learned as patterns by speakers. Therefore, the theoretical apparatus advanced here is not to be taken as so ambitious a theory as to account for all such sandhi systems, but only a direction in which the analysis for phonetically natural sandhi may be sought. A complete understanding of synchronic Chinese tone sandhi would require much more careful phonetic documentation, psycholinguistic testing, and phonological theorizing, and consequently awaits much further research.

References

Archangeli, Diana. 1991. Syllabification and prosodic templates in Yawelmani. *Natural Language and Linguistic Theory* 9:231-283.

Boersma, Paul 1998. *Functional Phonology: Formalizing the Interaction between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphic.

Briggs, Elinor. 1961. *Mitla Zapotec Grammar*. Instituto Linguistico de Verano, Mexico.

Broselow, Ellen. 1995. Skeletal positions and moras. *Handbook of Phonological Theory*, ed. by John A. Goldsmith, 175-205. Oxford: Blackwell.

Cahill, Michael. 1999. *Aspects of the Morphology and Phonology of Kɔnni*. Columbus: Ohio State University dissertation.

Clark, Mary M. 1983. On the distribution of contour tones. *Proceedings of WCCFL 2*, 44-55.

Clements, George N., and Kevin C. Ford. 1979. Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry* 10:95-108.

Cui, Rong-Chang. 1997. *Chengduhua Yindang* [*The Sound Record of Chengdu*], ed. by Jing-Yi Hou. Xiandai Hanyu Fangyan Yinku [The Sound Archives of Modern Chinese Dialects]. Shanghai: Shanghai Jiaoyu Chubanshe.

Doke, Clement M. 1937. An outline of ǂKhomani Bushman phonetics. *Bushmen of the Southern Kalahari*, ed. by J. D. Rheinallt Jones and C. M. Doke, 61-88. Johannesburg: Witwatersrand University Press.

Duanmu, San. 1990. *A Formal Study of Syllable, Tone, Stress and Domain in Chinese Languages*. Cambridge: MIT dissertation.

Duanmu, San. 1994a. Against contour tone units. *Linguistic Inquiry* 25:555-608.

Duanmu, San. 1994b. Syllabic weight and syllabic duration: A correlation between phonology and phonetics. *Phonology* 11:1-24.

Edmondson, T., and J. T. Bendor-Samuel. 1966. Tone patterns of Etung. *Journal of African Languages* 5.1:1-6.

Elwell-Sutton, Laurence Paul. 1976. *The Persian Metres*. Cambridge: Cambridge

University Press.
Everett, Daniel, and Keren Everett. 1984. On the relevance of syllable onsets to stress placement. *Linguistic Inquiry* 15:705-711.
Flemming, Edward. 1995. *Auditory Representations in Phonology*. Los Angeles: University of California dissertation.
Fu, Guotong. 1984. Wuyi fangyan de liandu biandiao [Tone sandhi in the Wuyi dialect]. *Fangyan* [*Dialects*] 1984.2:109-127.
Goldsmith, John. 1976. *Autosegmental Phonology*. Cambridge: MIT dissertation.
Gordon, Matthew. 1999. *Syllable Weight: Phonetics, Phonology, and Typology*. Los Angeles: University of California dissertation.
Harms, Robert T. 1962. *Estonian Grammar*. Indiana University Publications, Uralic and Altaic Series, Vol.12. The Hague: Mouton.
Hayes, Bruce. 1979. The rhythmic structure of Persian verse. *Edebiyat* 4:193-242.
Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistics Inquiry* 20:253-306.
Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
Hou, Jingyi. 1980. Pingyao fangyan de liandu biandiao [Tone sandhi in the Pingyao dialect]. *Fangyan* [*Dialects*] 1980.1:1-14.
Hyman, Larry M. 1985. *A Theory of Phonological Weight*. Dordrecht: Foris Publications.
Hyman, Larry M., and Francis X. Katamba. 1990. Final vowel shortening in Luganda. *Studies in African Linguistics* 21:1-59.
Hyman, Larry M., and Francis X. Katamba. 1993. A new approach to tone in Luganda. *Language* 69:34-67.
Keating, Patricia A. 1988a. The phonology-phonetics interface. *The Cambridge Survey*, Vol.1: *Linguistic Theory*, ed. by F. J. Newmeyer, 281-302. Cambridge: Cambridge University Press.
Keating, Patricia A. 1988b. Underspecification in phonetics. *Phonology* 5:275-292.
Keating, Patricia A., and Abigail Cohn. 1988. Cross-language effects of vowels on consonant onsets. *JASA* 84, Supplement 1: S84.
Kirchner, Robert. 1997. Contrastiveness and faithfulness. *Phonology* 14:83-111.
Lanham, L. W. 1958. The tonemes of Xhosa. *African Studies* 17:65-81.
Lanham, L. W. 1963. The tonemes of Xhosa: A restatement. *Studies in Linguistics* 17: 35-58.
Leather, Jonathan. 1983. Speaker normalization in perception of lexical tone. *Journal of Phonetics* 11:373-382.
Lehiste, Ilse. 1972. Timing of utterances and linguistic boundaries. *Journal of the*

Acoustical Society of America 51:2108-2124.
Lindblom, Björn, Bertil Lyberg, and Karin Holmgren. 1981. *Durational Patterns of Swedish Phonology: Do They Reflect Short-Term Motor Memory Processes?* Bloomington: Indiana University Linguistics Club.
Lyberg, Bertil. 1977. Some observations on the timing of Swedish utterances. *Journal of Phonetics* 5:49-59.
Manuel, Sharon Y. 1990. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *JASA* 88:1286-1298.
McCarthy, John, and Alan Prince. 1986. *Prosodic Morphology*. Manuscript. University of Massachusetts and Brandeis University.
McCawley, James D. 1968. *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.
Miller, J. L., and F. Grosjean. 1981. How the components of speaking rate influence perception of phonetic segments. *Journal of Experimental Psychology: Human Perception and Performance* 1:208-215.
Moore, Corinne B. 1995. *Speaker and Rate Normalization in the Perception of Lexical Tone by Mandarin and English Speakers*. Ithaca: Cornell University dissertation.
Moore, Corinne B., and Allard Jongman. 1997. Speaker normalization in the perception of Mandarin Chinese tones. *JASA* 102:1864-1877.
Mugele, Robert. 1982. *Tone and Ballistic Syllable in Lalana Chinantec*. Austin: University of Texas dissertation.
Newman, Paul. 1972. Syllable weight as a phonological variable. *Studies in African Linguistics* 3:301-323.
Ourso, Meterwa. 1989. *Lama Phonology and Morphology*. Champaign: University of Illinois, Urbana Champaign dissertation.
Paster, Mary. 1999. *Issues in the Tonology of Gã*. Senior honors thesis. Columbus: Ohio State University.
Pols, Louis C. 1986. Variation and interaction in speech. *Invariance and Variability in Speech Processes*, ed. by Joseph S. Perkell and Dennis H. Klatt, 140-162. Hillsdale, NJ: Lawrence Erlbaum Associates.
Port, Robert F. 1979. The influence of tempo on stop closure duration as a cue for voicing and place. *Journal of Phonetics* 7:45-56.
Pulleyblank, Douglas. 1986. *Tone in Lexical Phonology*. Dordrecht: D. Reidel Publishing Company, Kluwer Academic Publishers Group.
Snoxall, R. A. 1967. *Luganda-English Dictionary*. Oxford: Oxford University Press, Clarendon Press.
Snyman, Jannie Winston. 1970. *An Introduction to the !Xũ! (!Kung) Language*. Communications from the School of African Studies, University of Cape Town,

No.34. Cape Town: A. A. Balkema.
Steriade, Donca. 1991. Moras and other slots. *Proceedings of the 1st Meeting of the Formal Linguistics Society of the Midwest*, ed. by D. Meyer and S. Tomioka. Madison: University of Wisconsin.
Steriade, Donca. 1999. Phonetics in phonology: The case of laryngeal neutralization. *Papers in Phonology* 3, ed. by Matthew K. Gordon, 25-146. UCLA Working Papers in Linguistics, No.2.
Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. *Papers in Laboratory Phonology*, Vol.5: *Language Acquisition and the Lexicon*, ed. by Michael Broe and Janet Pierrehumbert, 313-334. Cambridge: Cambridge University Press.
Steriade, Donca. 2001. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. Manuscript. Los Angeles: University of California.
Stevick, Earl W. 1969. Pitch and duration in Ganda. *Journal of African Languages* 8:1-28.
Strangert, Eva. 1985. *Swedish Speech Rhythm in a Cross-language Perspective*. Acta Universitatis Umensis, Umeå Studies in the Humanities 69. Almqvist and Wiksell International, Stockholm, Sweden.
Sundberg, Johan. 1979. Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics* 7:71-79.
Tauli, Valter. 1973. *Standard Estonian Grammar*, Part 1: *Phonology, Morphology, Word-formation*. Uppsala: Almqvist & Wiksell.
Trubetzkoy, Nikolai S. 1939. *Grundzüge der Phonologie*. Güttingen: Vandenhoeck and Ruprecht. English translation by Christiane A. M. Baltaxe. Berkeley and Los Angeles: University of California Press, 1969.
Yip, Moira. 1989. Contour tones. *Phonology* 6:149-174.
Yip, Moira. 1995. Tone in East Asian languages. *Handbook of Theoretical Phonology*, ed. by John A. Goldsmith, 476-494. Oxford: Blackwell.
Yue-Hashimoto, Anne O. 1987. Tone sandhi across Chinese dialects. *Wang Li Memorial Volumes*, English Volume, ed. by Chinese Language Society of Hong Kong, 445-474. Hong Kong: Joint Publishing Co.
Zec, Draga. 1988. *Sonority Constraints on Prosodic Structure*. Stanford: Stanford University dissertation.
Zee, Eric, and Ian Maddieson. 1979. Tones and tone sandhi in Shanghai: Phonetic evidence and phonological analysis. *UCLA Working Papers in Phonetics* 45:93-129.
Zhang, Jie. 1998. Checked tones in Chinese dialects—Evidence for phonetically-driven phonology. *Proceedings of CLS 34*, Vol.1, 439-453.

Zhang, Jie. 1999. Duration in the tonal phonology of Pingyao Chinese. *Papers in Phonology* 3, ed. by Matthew K. Gordon, 147-206. UCLA Working Papers in Linguistics, No.2.

Zhang, Jie. 2000a. The phonetic basis for tonal melody mapping in Mende and elsewhere. *Proceedings of WCCFL 19*, ed. by Roger Billerey and Brook Lillehaugen, 603-616. Somerville: Cascadilla Press.

Zhang, Jie. 2000b. Non-contrastive features and categorical patterning in Chinese diminutive suffixation—Max[F] or Ident[F]? *Phonology* 17:427-478.

Zhang, Jie. 2001. The contrast-specificity of positional prominence—Evidence from diphthong distribution. Paper presented at the 75th Annual Meeting of the LSA. Washington, DC.

Zhang, Jie. 2002. *The Effects of Duration and Sonority on Contour Tone Distribution—A Typological Survey and Formal Analysis*. New York: Routledge.

Zhang, Jie. (in press). The role of contrast-specific and language-specific phonetics in contour tone distribution. *The Phonetic Basis of Phonology*, ed. by Bruce Hayes, Robert Kirchner, and Donca Steriade. Cambridge: Cambridge University Press.

Zhang, Shengyu. 1979. Chaoyang fangyan de liandu biandiao [Tone sandhi in the Chaoyang dialect]. *Fangyan* [*Dialects*] 1979.2:93-121.

Zhang, Shengyu. 1980. Chaoyang fangyan de liandu biandiao II [Tone sandhi in the Chaoyang dialect II]. *Fangyan* [*Dialects*] 1980.2:123-136.

[Received 7 March 2003; revised 15 June 2004; accepted 19 June 2004]

Department of Linguistics
University of Kansas
1541 Lilac Lane
Blake Hall, Room 427
Lawrence, KS 66045
USA
zhang@ku.edu

起伏調認可與起伏調表式

張　杰

堪薩斯大學

我在本文中將論證：起伏調位置上的限制，並非來自雙音拍的認可要求。這些論證包括：不以雙音拍爲優先的音節位置在認可上的優勢；起伏調標記區別的額外層次，這些額外層次超過了理論所允許的最大音拍數；帶相同數目音高目標的不同起伏調之間的差異；長期存在的音拍不一致的問題。我將提出一種理論上的設計，允許調型和韻長的語音細節得以出現在音韻表式之中。這種設計一方面可以對跨語言的起伏調認可行爲提供較佳的說明，一方面也不致削弱音韻理論的解釋力。

關鍵詞：起伏調，音拍，音長，音韻語音介面