13. Tone in East Asian Languages

MOIRA YIP

Subject	Theoretical Linguistics » Pholonogy
Place	Asia » Eastern Asia
DOI:	10.1111/b.9780631201267.1996.00015.x

0 Introduction

Theoretical phonologists have devoted more attention to African tone languages than to East Asian ones, probably because phonologists thrive on alternations, and the morphological characteristics of East Asian languages make alternations (segmental or tonal) relatively rare. Given this, one might ask what we can learn about tone from East Asian languages that we cannot learn from African languages. One striking observation about East Asian languages is the richness of their tonal inventory. Where African languages typically contrast two, perhaps three level tones (see chap. 12, this volume), East Asian languages, particularly Chinese languages, frequently contrast four levels and several rising or falling (i.e., contour) tones. They are thus fertile territory for exploring the distinctive features of tone – what they are, how they realte to segmental features, and how they are organized in the feature geometry. The first and longest section of this chapter is devoted to this question.

A second way in which East Asian languages may be helpful is in illuminating our understanding of what may be the tone-bearing unit (TBU) in language. Some Chinese languages appear to use the mora, so that only long syllables may bear two tones, which surface as contours. However, I will adopt the view that the lack of contour tones on short syllables is a phonetic effect, and any syllable may phonologically bear any tone. It then follows that the TBU is the syllable, not the mora. I will then discuss the metrical foot, which typically limits tones to one per foot (Mandarin); this raises the issue of the relationship between tone and stress, which can coexist in Chinese. Finally I discuss the prosodic word and the phonological phrase (Shanghai).

A third area of investigation addresses the types of tone sandhi rules that we find. What form do the rule themselves take (assimilation, dissimilation, rules conditioned by phrasal position, etc.)? The remaining issues will receive only brief mentions for reasons of space. One observation that needs an explanation is the apparent absence of downstep in East Asian languages.

Another cluster of issues arises over how to define the domains within which sandhi rules apply.

In a summary like this, each of these issues receives only a cursory treatment; it should be clear that there are many questions awaiting further research. I give more attention to Chinese languages than other East Asian language families as an unavoidable consequence of the fact that I am more comfortable making generalizations about Chinese that I am about such other East Asian tone languages as those of Tibeto-Burman, Miao-Yao, Tai, and Austroasiatic (Vietnamese). I draw heavily on the work of previous researchers, especially that of Matthew Chen, Bao Zhi-ming, and Duanmu San, without whose insights this chapter could not have been written. Since this article is for the general phonologist, rather than the specialist in Chinese, I have largely limited the references to English-language sources, even though these are often secondary works using data from Chinese-language field studies. Specialist readers may pursue the original sources for themselves.

1 Tonal Features

1.1 Six Proposals for Tonal features

A satisfactory feature system for tone must meet the familiar criteria of characterizing all and only the contrasts of natural language, the appropriate natural classes, and allowing for a natural statement of phonological rules and historical change. In looking at East Asian tone systems the main issues are these:

(1) (a) How many different levels must be represented?

(b) Are contour tones single units or sequences of level tones?

(c) What is the relationship between tonal features and other features, especially laryngeal features?

These issues are discussed in great detail by many researchers, most notably Anderson (1978), Yip (1980, 1989), Pulleyblank (1986), Hyman (1986, 1989), Snider (1990), Bao (1990), Duanmu (1990a), and references cited therein. The features I will use are Register, a binary feature referred to in Yip (1980) as [+/ -upper], but which I will here call upper case H/L for convenience, and Pitch (following the usage of Duanmu 1990a), also a binary feature, denoted here by lower case h/l. Two binary features can produce four combinations, giving four level tones. Hyman and Duanmu differ from the other authors cited above in that they allow for a three-way contrast in both register and pitch. The resulting nine levels are not attested in any language, so I continue to restrict the system to a four-way contrast. The languages that have been described as having five level tones (notably Black Miao (Chang 1953), see Duanmu for a good summary) may require revision of this assumption, but so little is known about the phonological behavior of these tone systems that some of the "level" tones may turn out to be phonologically contours.

Contour tones (rising or falling tones) are sequences of Pitch values, with a single Register value. For example, [H, lh] is a high rising tone, whereas [L, hl] is a low falling tone. We therefore get the following eight-tone inventory:

(2)

H, h	L, h
H, l	L, l
H, hl	L, hl
H, lh	L, lh

This system is found almost perfectly in Cantonese; the digits are the Chao "tone letter" system, in which 5 denotes high pitch and 1 low pitch, used by all workers on Chinese tone; a single digit is used to denote an obstruent-final (i.e., short) syllable.

(3)

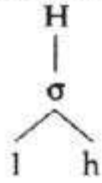
H, h	55	si	"poem"	L, h	22	si	"affair"
H, l	33	si	"try"	L, l	—		
H, hl	53	si	"silk"	L, hl	21[1]	si	"time"
H, lh	35	si	"cause"	L, lh	24	si	"city"

Morphemes are almost exclusively monosyllabic in Chinese, and with rare exceptions each lexical morpheme has one of the lexical tones. In some dialects some suffixes may be toneless.

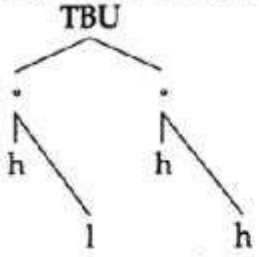
Since the work of Clements (1985), Sagey (1986), and others on segmental feature geometry (see chap. 7, this volume), the arrangement of the tonal features has been an object of inquiry. A variety of different models have been proposed, each making slightly different predictions. One of the fundamental ideas of feature geometry theory is that groups of features that spread together are grouped together under a single node. I will use spreading as the diagnostic for constituency in the feature geometry, and survey the extant proposals for tonal features, and the evidence. The major proposals of which I am aware are given below; I will not discuss earlier proposals such as Wang (1967) or Woo (1969) because their conversion into autosegmental terms is not obvious. I have grouped the proposals by the structural relationship between the Register and Pitch features (independence, sisterhood, and dominance) and standardized the terminology across proposals as much as possible. Proposals differ in what they take to be the tone-bearing unit (TBU), and these differences are retained here. I show a high rising tone, [35], in each system:[2]

(4)

(a) Independence (Yip 1980b)

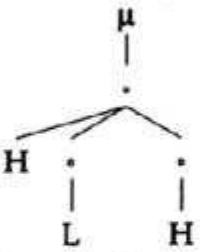


(b) Sister-hood (Bao 1990b; Duanmu 1990a; Clements 1989a; Hyman 1989b; Snider 1990b) Clements, Snider:[3]



Tonal Nodes

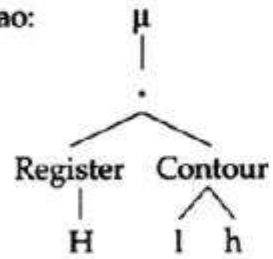
Hyman: (extrapolating from his fn. 3 on East Asian contours)



Tonal Root Node

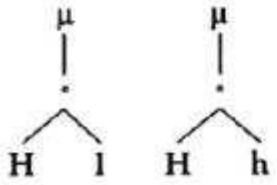
Tonal Node

Bao:



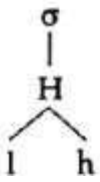
Tonal Node

Duanmu:



Laryngeal Node

(c) Dominance (Yip 1989a)



These theories make the following predictions. (Note that all theories allow terminals to spread):

(5)

	Spreading		
	Whole Tone	Register only	Contour as a whole, w/out register
Yip 1980b		x	
Clements		x	
Hyman	x	x	
Yip 1989a	x		
Bao	x	x	x
Duanmu		x	

I will briefly review the evidence for each type of spreading. I will start with contour spreading as a whole, but without register.

1.2Contour Spreading

Bao's is the only model which allows for spreading contour without register. It seems, however, that the cases he offers as instances of such spreading are not entirely convincing, and I will conclude that in the absence of more clearcut cases a more restrictive model is to be preferred.

Bao (1990b, p. 96ff.) offers Zhenjiang and Wenzhou as cases of spreading just a contour node. Zhenjiang (Jiangsu province, Zhang 1985) is analyzed by Bao as spreading the property of being level (i.e., the contour node) from a final syllable back onto a penultimate syllable. Data from Zhenjiang are given below; these changes apply to the penultimate syllable of domains of two to three syllables long. Zhang tells us that the initial syllable in three syllabled cases is always level, with /55,5/ staying unchanged and all others becoming 33.

(6)

$\sigma_1 \backslash \sigma_2$	42	31	35	55	5
42				33	
31		35			
35				22	
55			55		
5			5		

The leveling is found for falling tones not just before level tones, but before the rising 35 tone as well. This throws doubt on Bao's analysis, which would have to treat these cases separately. The general character of the system appears to involve greater neutralization and leveling the further from the end of the word the syllable lies. Antepenultimate syllables are all level, penultimate syllables are level or rising, and only final syllables may fall. I thus conclude that Zhenjiang is not a clear case of contour spreading.

Another reason for finding Zhenjiang less than wholly convincing is that spreading of levelness will never be as convincing as spreading of a rise or a fall, since one cannot distinguish spreading of the contour node from spreading of the subordinate tone. Wenzhou (Zhengzhang 1964) is more interesting since Bao claims that /hl/ spreads as a unit, but without register. Underlining indicates glottalization. Bisyllabic compounds composed of (1) 45, 34, or 22 before falling 42, 31, or (2) 42,

323, or 212 before falling 42 surface as falling-falling [42-21]. Bao analyzes this as the spread of the /hl/ contour from the final syllable backwards. There are empirical problems with the analysis, however, when extended to the larger data set in Bao's appendix. Bao predicts that all phrases ending in an underlying /hl/ should show this spread, but they do not. For example, all tones except 45, 34, and 22 before 31 surface as 22-2, level throughout both syllables. Conversely, we find falling syllables before the underlyingly level /22/, and even before the underlying rise, 34. I should also note that Duanmu (p. 147) argues against Bao's analysis on several other grounds, and points out that in the 42-21 pattern the [21] may be phonologically level, and indeed has a [42-1] alternant.

In order to understand what is going on, we again need to look at the complete tonal system of Wenzhou. Wenzhou is a Wu dialect, and it has eight contrasting tones on monosyllables. On longer phrases there are still only eight possible patterns, determined by the underlying tone of the last, and sometimes also the penultimate, syllable. On spans of three, four, or five syllables, it is clear that if the last two syllables are one of the sequences analyzed by Bao as spreading a /hl/ contour to give 42-21, the pattern of the whole phrase is a consistent exception to the general pattern of Wenzhou. The data is given below. The letter *t* stands for tones whose value depends on the underlying tones of the final two syllables in a non-obvious way. The phrases of interest here are those in the lefthand column; note that it is not only the final syllables themselves that are different, but the *preceding* span as well.

(7)

	Final two syllables 42–21 in [σσ] cases	All other phrases
σσσ	42–21–21 / 21–42–21	43–t-t
σσσσ	42–21–31–21	34–43–t-t / 2–4–t-t
σσσσσ	42–21–31–21–21	2–4–43–t-t

Bao and Duanmu both note that stress plays a role in Wenzhou sandhi, since 42-21 phrases have initial stress, and there is a 21-42 variant with final stress. In longer spans there is an alternating tonal pattern, with a 31 breaking the string of 21s. These patterns may in fact be the result of imposing alternating stress on a toneless span, with no tone sandhi involved at all. In this view, 42-21-31-21-21 is simply the phonetic instantiation of a stress grid:

(8)

As to why the tone deletes in these particular circumstances, I have no answer at present, but I conclude that Wenzhou is also not a clear case of spreading contour alone.

Since neither of the putative cases of contour node spreading seems to hold up under closer scrutiny, I conclude for now that our tonal model should not admit this possibility, and opt for one of the more restrictive models. I now turn to whole tone spreading.

1.3Whole Tone Spreading

Hyman, Yip (1989) and Bao allow spreading of a complete tone. Bao offers two examples of whole tone spreading, Changzhi and Danyang. In Changzhi (Shanxi province; Bao 1990, using data from

Hou 1983) we have the following pattern upon attachment of the suffixes /ti/ or /tə?/ to a base:

(9)

$suaŋ^{213}$	$suaŋ^{213}$ ti^{213}	"sour"
$xuaŋ^{24}$	$xuaŋ^{24}$ ti^{24}	"yellow"
$yaŋ^{535}$	$yaŋ^{535}$ ti^{535}	"soft"
$laŋ^{53}$	$laŋ^{53}$ ti^{53}	"rotten"

Bao analyzes this as whole tone spreading, but Duanmu (p. 145) suggests that Changzhi reduplicates the tone rather than spreading it. However, Hou's data shows quite clearly that Bao is right, and this must be whole tone spreading. The evidence comes from overt reduplication, which shows quite different tonal patterns, with sandhi rules that dissimilate one or another of the reduplicated tones.

(10)

/213/	$saŋ^{213}$ $saŋ^{35}$	"fan"
/53/	$tuŋ^{35}$ $tuŋ^{53}$	"move"?
/535/	$ts'ɔ^{535}$ $ts'ɔ^{35}$	"fry"
/24/	$tɕ'iəu^{24}$ $tɕ'iəu^{53}$	"ask for, beg"

Although the reduplication is obscured by the subsequent sandhi rules, note that neither syllable has a fixed tone, showing that tonal reduplication must have taken place. If (10) is the result of tonal reduplication, the data in (9) must be something else, and spreading is the obvious candidate. I conclude that Changzhi shows true spreading of an entire tone.

The last case to be discussed here is Danyang, where one phrasal tone pattern is 42–42–42–…24, roughly conditioned by an initial /24/ tone. Using data from Lü (1980), Chen (1986), Chan (1988), Yip (1989), and Bao have proposed analyses that in one way or another involve spreading of the entire contour tone, usually accompanied by a sandhi rule that changes 24 to 42 before another 24, although details vary. There is some question, however, as to whether this is spreading of any kind. Duanmu argues that the 42–42–42…24 pattern is derived from a sequence of underlying /24/ tones, /24–24–24…24/, followed by the sandhi rule that changes 24 → 42 / → 24, and no spreading is involved. Wang Hong-jun (1991)shows clearly that this is not the case, and that some of these patterns do not have underlying /24/ throughout. What determines the phrasal pattern is the tone of the first syllable alone, provided its historical origins are taken into account (specifically, it is phrases that begin with a Ze, a literary style syllable, and we must assume that just as English still distinguishes between the Greco-Latin and the native vocabulary, or Malayalam between Dravidian and Sanskrit vocabulary, Danyang distinguishes Ping and Ze, and literary and colloquial). It would seem, then, that Danyang can be maintained as a case of a complete contour tone spreading as a unit, although it is not as clear-cut as the Changzhi example. The only two models which allow this, without also allowing contour node spreading without Register, are those of Hyman (1989) and Yip (1989). They differ in whether they allow independent Register spreading, a matter to which I now turn.

1.4Register Spreading

Bao gives two cases that he analyzes as spreading just Register. In Wuyi, a Wu dialect, two-syllable nominal compounds with the first syllable /24/, high rising, and the second syllable either /24, 213, 53, or 31/, surface with the second syllable high falling, [53]. Data from Bao (p. 209), citing Fu (1984).

(11)

/24–24/	t'ie koŋ	[24–53]		"Heavenly Lord"
/24–213/	hua fioŋ	[24–53]	[hua ʔoŋ]	"a fruit"
/24–31/	sa vuo	[24–53]	[sa fuo]	"raw meal"

Bao and Duanmu both analyze this as spreading of H Register, since the rising contour does not appear to spread too. However, since the final syllable is always falling, it is not possible to tell whether (1) the whole tone with its contour, or (2) just the Register, has been spread, since neutralization to falling on the final syllable (possibly by addition of a final low tone) is needed in both accounts.[4] Since the previous section demonstrated the need for whole tone spreading, and the Wuyi data can be dealt with by this means, Wuyi cannot be used to establish the need for Register spreading.

The other case offered as Register spreading is Pingyao (Hou 1980). Schematic Pingyao facts are given in (12), with the sequences of most interest in boldface:

(12)

σ_1	13	35	53
13	13–13	31–35	**35–423**
35	**13–13**	31–35	35–423
53	53–13	53–35	35–423

13 → 35 / ____ 53
35 → 13 / ____ 13

For Bao, rising tones receive Register from the following syllable, so that in the first column the /35/ initial syllables acquire L Register from the following /13/, and in the third column /13/ becomes /35/ before the H Register /53/. In the second column, an earlier metathesis rule changes the first syllable /lh/ to /hl/, so Register spread never applies. Chen (1991) gives an alternative analysis of Pingyao in which Register does not spread, but is neutralized under the influence of the Pitch features of the following tone, so that for rising tones Register becomes H before hl (column 3), and L before lh (column 1). He retains Bao's metathesis rule. A somewhat similar connection between pitch and Register features has been observed in Hausa, by Inkelas, Leben and Cobler (1987), although I know of no other cases in Chinese. Pingyao remains the most convincing case of Register spread I have encountered, but its restricted nature (only rising tones are targets, and it is bled by a metathesis rule) and the alternative offered by Chen's analysis makes it less than totally convincing.[5] I conclude that in our present state of knowledge the only kinds of spreading for which we have strong motivation are spreading of terminal features, and spreading of the entire tone, and we should adopt the most restrictive model compatible with these types of spreading. The only model in (5) with these properties is one where Register dominates the pitch features, and there is no special contour node, as in Yip (1989), and for now I will use this as my working model.

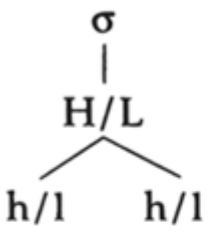
1.5Tonal and Laryngeal Features

A brief note is needed here on a major topic: the relationship of tonal and other laryngeal features. The earliest attempt to characterize a set of features that would relate tone, voicing, aspiration, and glottalization was that of Halle and Stevens (1971). Both Bao and Duanmu take the strong position that identifies Register and obstruent voicing as one and the same thing. Bao allows a two-way distinction, using the feature [+/-stiff vocal cords]. Duanmu adopts the same idea, but uses the pair of features [stiff vocal cords] and [slack vocal cords] for a three-way distinction. Putting the details aside, if we equate Register and voicing, and incorporate this insight into the working model of Yip

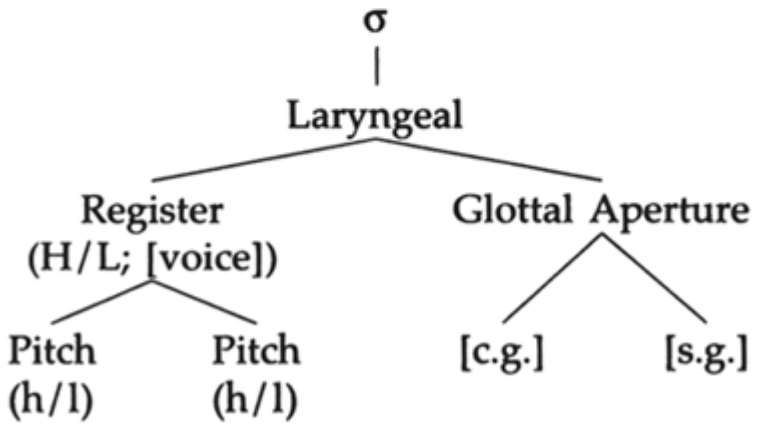
(1989) arrived at in the previous sections, the result is shown in (13). (13a) shows the purely tonal model; (13b) substitutes a Register node, variously realized as H/L tonal register or [voice], for the plain H Register, and a Pitch node, realized as h/l Pitch, for the simple h/l Pitch features. This Pitch node appears twice, to allow for contour tones. The Glottal Aperture node dominates the features Constricted Glottis (c.g.) and Spread Glottis (s.g.). I will not discuss the issue of whether the Laryngeal node associates directly to the syllable, as shown here, or via the segmental Root node.

(13)

(a) Yip (1989)



(b) Incorporating this into a laryngeal feature model:



Duanmu makes two claims: first, languages with too many tonal levels to deal with by the Pitch distinction alone, and that thus use contrastive Register, will always show voice quality distinctions alongside pitch distinctions. This claim is probably too strong. Cantonese, for example, which makes a rich use of Register, as shown in (3) above, is not reported to have voice quality distinctions between the Registers.

The second consequence is that spreading of Register or the whole tone should, ceteris paribus, interact with the voicing of intervening obstruents. Obstruents should block the spreading of incompatible Register values, or assimilate in voicing to the spreading value. The potential blocking effect means that any spreading involving Register will be rare, as Duanmu points out. The rarity of whole tone spreading (the only way Register can spread in the view of this paper) is striking. In Wuyi, where high registered tones spread, the spreading is indeed accompanied by onset devoicing, as shown by the data from (11), repeated here:

(14)

/24–24/	t'ie koŋ	[24–53]		"Heavenly Lord"
/24–213/	hua fioŋ	[24–53]	[hua ʔoŋ]	"a fruit"
/24–31/	sa vuo	[24–53]	[sa fuo]	"raw meal"

In Changzhi, on the other hand, whole tone spreading does not have any effect on obstruent voicing, even when a low register tone spreads across a voiceless obstruent, as in [suan^{213} ti^{213}] "sour". Since Changzhi (unlike Wuyi) has no voiced obstruents, this may perhaps be attributed to structure preservation.

The intimate connection between tone and laryngeal features is very clear when one looks at the historical development of tone (see Kingston and Solnit

1988 for a detailed study). In some dialects today this clear connection can still be seen. The Wujiang dialect (Ye 1983), brought to my attention by Ting Pang-Hsin, has twelve tones, which would seem to exceed the limits of the system. Arranged according to laryngeal qualities of the onset, and taking the short (Ru sheng) tones as variants of the level tones, however, they are reducible to the more manageable three:

(15)

	h	hl	hlh
Voiceless plain, including fricatives, onsetless, and /h/.	55/5	51	412
Voiceless aspirated	33/3	42	312
Voiced, obstruent, and sonorant, including	13/2	31	212

Following Duanmu, and equating voice with low register, the voiced series in (15) will be L Register, and the others H. The difference between the plain and aspirated sets will be in [+/-spread glottis], so only six tones will contrast underlyingly (where voicing and low register are one and the same feature). Why aspiration should have a lowering effect is unclear, and for detailed discussion of this point I refer the reader to Kingston and Solnit's paper.

The two most noticeable interactions between laryngeal features and tone are the effects of voiced obstruent onsets, touched on above, and the effects of glottal codas, and I will discuss each in turn. Voiced obstruent onsets usually lower pitch, as in Wujiang (above), Shanghai, or Tibetan. Tibeto-Burman languages offer fertile ground for the tone-larynx connection. In these languages there is an interaction between onset voicing, voice quality on the nuclear vowel, and tone. Consider Eastern Tamang as described by Weidert (1987, p. 260ff.). There are four "tones," two of which are associated with clear phonation and two with breathy phonation. Voiced obstruents and breathy phonation go together, as do voiceless obstruents (plain or aspirated) and clear phonation. The two breathy tones are lower in pitch than the two clear tones. Within a single phonation type, the tones are distinguished by level or shape:

(16)

Voiceless obstruent onsets:	Clear phonation:	H (or HL) M
Voiced obstruents:	Breathy phonation:	LM L

Since the four "tones" are contrastive in sonorant-inital syllables, they can be taken as phonemic and obstruent voicing as derived. If we follow Bao and Duanmu in identifying low register with obstruent voicing, this is a straightforward assimilation rule, spreading low register/obstruent voice leftwards. Such cases make clear the need for the feature geometry to capture the close relationship between tonal and laryngeal features. Tibetan shows interesting neutralizations in compounding that lead us

to the same conclusion.

See Yip (1990) and Meredith (1990) for two different recent accounts of the Tibetan facts. Within Chinese the Wu dialects show similar phenomena; see Bao (1990) and Duanmu (1990a), and references cited therein. Also Duanmu (1990b) on Shanghai.

The full picture of the tonal-laryngeal relationship is undoubtedly more complex than I have implied so far. Burmese (Bradley 1981) has a three-way contrast in phonation type and pitch between plain (clear) (L), breathy (M/ HM) and creaky (HL). The tones are my interpretion of Bradley's description. Although Burmese is a Tibeto-Burman language, this system of voice quality distinctions is typical of that found in Mon-Khmer languages. Exactly how to translate this into tonal and laryngeal features is quite unclear, and I leave this for further research. See Duanmu (1990a, p. 125ff.) for one approach.

An interesting problem arises if we take the model to which we have been led by the East Asian data, and apply it to African languages. Typically H and L tones spread freely across long spans of syllables, across any kind of consonant (depressor consonants aside). For example, Luganda (Hyman, Katamba, and Walusimbi 1987) has a HL pattern on accented words, and the H span may encompass voiced stops, the low portion may encompass voiceless stops.[6] If Duanmu is right about identifying voice and register, the spreading in African languages, which does not usually interact with voicing, must be Pitch, not Register, and yet in my model Pitch is dependent on Register, and can thus only be specified if Register is first specified. But if the language only distinguishes two surface tones, surely only Pitch *or* Register is being used distinctively, not both. The only resolution to this impasse that occurs to me is that the dependency relations between Register and Pitch might differ between Chinese and African languages, along the lines discussed by Mester 1986 and Selkirk 1988 for vowel and consonant features).

The second widespread laryngeal-tonal connection is the effect of glottal codas and unreleased (perhaps glottalized) stop codas. These typically shorten the syllable, and often permit only level tones on the preceding vowel. Cantonese, for example, contrasts seven tones on vowel- and nasal-final syllables, but allows only the three level tones on obstruent-final syllables. A particularly interesting case of interaction between glottal stops and tone is reported in Weidert (1987, p. 420). Boro is a Tibeto-Burman language with no tone contrasts on monosyllables. In bisyllables, however, a tonal contrast is observed on the second syllable; the acute accent marks high pitch, and unaccented vowels are mid.

(17)

dɣisaʔ	"rivulet"	daosáʔ	"chicken"
bajiʔ	"fifty"	dojíʔ	"sixty"

It turns out that this pitch difference is predictable from the presence or absence of a final glottal stop in the underlying form of the first morpheme. The glottal stop is lost in compounding, but surfaces as high tone on the following syllable.

(18)

dɣi	"water"	dɣisaʔ	"rivulet"
daoʔ	"fowl"	daosáʔ	"chicken"
ba	"five"	bajiʔ	"fifty"
doʔ	"six"	dojíʔ	"sixty"
jiʔ	"ten"		

No current theory of tonal features can explain the details of this process, but it serves to illustrate the intimate relationship between the various tonal and laryngeal features.

2 The Tone-bearing Unit

2.1Obstruent-Final Syllables

In some African languages (e.g., Luganda, Clements 1986), contour tones appear only on long vowels or closed syllables. This suggests that the TBU is the mora, and tone association is strictly one-to-one (though see chap. 12, this volume). In Chinese languages we observe a related but interestingly different situation. Sonorant-final syllables (vowel or nasal final) may bear any tone, but if a language has obstruent-final (i.e., unreleased plain voiceless stops) syllables, they may bear only a subset of tones, usually level tones. These syllables are shorter than the sonorant-final syllables (see Kao 1971). These facts can be explained given the following assumptions:

(19) Contour tones have two tonal root nodes.

The mora is the TBU.

Only sonorant codas can be moraic.

There is open-syllable lengthening.

The obstruent codas will be nonmoraic, and thus not TBUs, so such syllables will have only one TBU and hence only one (level) tone. Duanmu (1990a) takes a rather different approach. He argues that all Chinese syllables are bimoraic, so that even obstruent codas are TBUs.[7] The failure of contours to surface on these syllables is seen as phonetic: on an unreleased stop no tone can be realized, even though it is phonologically present. He shows that in Shanghai and Taiwanese (pp. 137–138) these tones may surface if the stop coda is deleted.

What both these analyses share is the assumption that contour tones are made up of two tonal root nodes, contra the model argued for in section 1. It is worth considering whether the lack of contour tones on stop-final syllables can possibly be explained if they are in fact one tonal Root Node, as argued for above. In fact, there is a straightforward way to understand these facts: I shall adopt Duanmu's view that phonetic realization is the reason that contour tones fail to surface on unreleased stop-final syllables, even though they may be phonologically present.[8] This move allows us to maintain the model in which contour tones are units (and thus explain why they can spread as a unit, as in Changzhi), while understanding their restricted distribution. This suggestion also removes the motivation for taking the mora to be the TBU in Chinese. If contour tones are single tonal nodes, and the syllable is the TBU, then all the facts follow. In fact, if the mora is the TBU, a contour tone could be associated to each mora, giving highly complex tones such as rise-rise, which are not found, so it is not only possible but essential that the syllable be the TBU in Chinese.[9]

2.2Distribution of Contour Tones in East Asian Languages

I would like to address the striking contrast between the distribution of contour tones in Chinese (where they are found on any syllable) and African languages (where they are found only at margins). There are three possible explanations for the typological difference. The first is that of Yip (1989), where I argue that in African languages underlying tones are all level, and contour tones are always the result of associating two tonal Root Nodes to one TBU, a configuration usually only possible at domain edges. In East Asian languages, by contrast, underlying tones include contours under a single Root Node, and these can associate to any single TBU just as a level tone can. This explanation best fits the feature model espoused here.

The second possible explanation is offered by Duanmu (1990a), for whom the facts follow from his assumption that only bimoraic syllables may bear contour tones, plus the claim that in Chinese *all* syllables are bimoraic, whereas in most African languages syllables are usually monomoraic, except when word-final lengthening provides an additional mora for a final contour tone.[10]

The last possible explanation is morphological. The best-studied group of African languages, the Bantu languages, has a rich affixing morphology producing long polysyllabic words, and roots do not usually surface without affixes. Roots and affixes may have tones or be toneless, and tone association takes place at the stem level, where stem consists of a root plus a number of suffixes (see Myers 1987 and Hewitt and Prince 1989 for details). Chinese languages, as is well-known, have little in the way of

affixes. Each lexical morpheme is a single syllable with its own tone, and can usually surface unadorned, although a prolific use of compounding means that most words are in fact bisyllabic or longer. If tone association is assumed to take place at the stem level as in Bantu, then in Chinese this means that each syllable will undergo tone association as a separate domain, (the few suffixes, all toneless, will be adjoined to the stem later). As a consequence each syllable will be the end of a domain for tone association in Chinese, and thus can bear a contour tone, whereas in Bantu only the last syllable of the relevant domain can bear a contour.

2.3Foot and Phonological Word

Before closing this discussion of tone-bearing units, let's look briefly at the behavior of tones in units larger than the syllable. In some dialects the stress foot plays a clear role in tonal phenomena. First, toneless syllables typically cannot bear stress, and adjoin to the preceding toned and stressed syllable to form a foot. Mandarin, Taiwanese, and many other dialects show this behavior. It is almost as if the foot is the TBU in such languages, and the tone surfaces on the head of the foot, i.e., the stressed syllable.

Second, the foot may be the domain of tone sandhi rules. Shih (1985) and Chen (1990) have argued that this is the case for the third-tone sandhi rule in Mandarin, but it is not entirely clear that “foot” is the right characterization of the prosodic unit involved, since this “foot” groups together two fully-toned syllables of roughly equal prominence, and cannot be identified with the other Mandarin foot created by grouping together one toned and stressed syllable with one or more toneless stressless syllables. Since in stress theory a foot means a grouping of one prominent syllable and one (or more, perhaps) less prominent syllable, I prefer to reserve the term for the units containing only one toned syllable. In that case the domain of third-tone sandhi must be larger, perhaps the prosodic word.

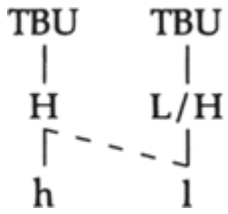
Leaving this open, another set of cases occurs in which the prosodic word more obviously plays a role. Selkirk (1986), Chen (1987), and Selkirk and Shen (1990) build prosodic words and phrases from syntactic information, and these then define tone sandhi domains in Amoy and Shanghai. In Amoy (Southern Min), which is sometimes reported to have final stress, nonfinal syllables change into specific sandhi tones. In Shanghai (Wu), which has initial stress, noninitial syllables lose their tones altogether, and the tone of the initial syllable controls the tone of the whole domain. Extending the notion of TBU still further, one could speculate that this is an example of the phonological phrase as TBU. In any case, there is a clear connection between the inability of toneless syllables to bear stress in Mandarin, and the inability of stressless syllables to bear stress in Mandarin, and the inability of stressless syllables to bear tone in Shanghai. Taking a wider view, these are in turn related to the known tendency of prominent syllables to attract tone in Bantu languages.[11] Meredith (1990) argues that in Mandarin stress is assigned on the basis of tone, with the high and high falling tones tending to attract stress, and the low tone tending to reject it. However, native speakers tend to reject the idea that there is any detectable stress in such words, and data collected in the phonetics lab at the Academy of Social Sciences in Beijing found no consistent judgements on the relative stress of fully toned syllables (Wang Hong-jun, p.c.).

3 Types of Sandhi

The types of tone sandhi found in Chinese are by and large familiar enough, at least in the clearly understood cases. There is a fundamental division between tone changes caused by a specific tonal environment, such as assimilation and dissimilation, and tone changes caused by purely positional factors, such as special phrase-final variants, or spreading over a phrasal domain.

As one would expect, among the first type we find both assimilation and dissimilation. Assimilation is common, as in Gao'an (Bao 1990a, p. 111). High tones / 55 / [H,h] become falling before any / I / tone, 33 [H,I] or 11 [L,I]. This is the result of the / 1 / spreading leftward onto the preceding syllable:

(20)



Dissimilation is also found, as in Tianjin. Data are from Chen (1985), the analysis is that of Yip (1989). Tianjin has four tones: 21, 45, 213, 53. When two tones are juxtaposed the following changes take place. All other sequences are unchanged.

(21)

Tones:	21 [L,l]	213 [L,lh]	45 [H,h]	53 [H,hl]
21 21	→ 213 21	l.l	→ lh.l	
213 213	→ 45 213	lh.lh	→ h.lh	
53 53	→ 21 53	hl.hl	→ l.hl	
53 21	→ 45 21	hl.l	→ h.l	

This is clearly dissimilatory, presumably OCP triggered, and, crucially, treats contours as units, as expected in our working model. Register plays no role here, and I assume that the accompanying register changes are secondary; that is, when 53 [H,hl] becomes 21 [L,l], it is the tones that dissimilate, and the Register change follows because in Tianjin all [l] tones are predictably [L].

Among the positional type of sandhi, the Wu dialects, like Shanghai, spread the tone of the first syllable over a phrasal domain (data from Xu, Tang, and Qian 1981–83). Here the tone of the first syllable spreads over the first two syllables.

(22)

Tone of first syllable	Tone of polysllabic domain
53	55 33 33 . . . 33 31
24	33 55 33 . . . 33 31
13	22 55 33 . . . 33 31

This pattern is easier to explain if contour tones have two tonal Root Nodes, and each then associates to one syllable of the longer domain. If contours are single Root Nodes, as proposed above, this sandhi may require some kind of splitting operation, obscuring its naturalness. Perhaps Shanghai has developed an African-style system, in which there are no underlying unitary contours.

Less well understood is the system common in the Min dialects, such as Taiwanese. Known as the Min tone circle, this was first discussed in a generative framework by Wang (1967). Each tone has two variants, one found in phrase-final position and one elsewhere. The details are given below:

(23)

Phrase-finally	55	35	53	11	33	2	4
Elsewhere	33	33	55	53	11	4/53	2/11

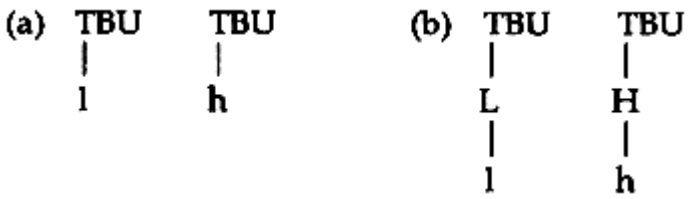
Attempts to write a single simple rule relating these two sets have largely failed; even though the analyses worked mechanically (notably Wang 1967; see also Yip 1980 for an unconvincing account), they failed to throw light on the phenomenon. Traditional scholars have viewed this as a case of paradigmatic replacement, but within generative phonology this is unappealing. Two recent attempts (Tsay 1990 and Truckenbrodt 1991) look at the data from the point of view of underspecification theory, and although each has some problems they come closer to a solution than anything else I have seen.

4 Additional Issues

Inevitably in a short article of this type there are areas that cannot be addressed; I mention a few here. Note that whether or not some topic has been covered in any detail or left till this section has less to do with the importance of the topic than with whether it seemed possible to do the issues justice in so brief an article.

One interesting question is why African languages often have downstep, and East Asian languages apparently do not. Downstep is a process whereby a low tone (which may or may not surface) causes lowering of a following high tone. One possibility is that the tonal geometry of the two language areas differs in such a way that two tones are adjacent in African languages but not in Asian languages. Specifically, suppose African languages have Pitch directly associated to the TBU, with Register either absent, or dependent on Pitch. This is shown in (24a). In Chinese-type languages, shown in (24b), Pitch features are lower in the feature tree, and an appropriate definition of locality might then block downstep effects in (b) type cases.

(24)



A major topic of current research on tone deals with how the domains within which tone sandhi rules apply can provide clues to the higher level prosodic structure of the string. Chen (1987, 1990), Selkirk and Shen (1990), Shih (1985), and others have investigated this complex issue. Some recent results are summarized in Chen (1990).

5 Conclusion

I have tried to show that the study of East Asian languages can tell us much about tone. On the assumption that distinctive features are part of Universal Grammar, the features required for Chinese must be those used by all tone languages, although of course, as with segmental features, no language will necessarily use all features distinctively. Given underspecification theory, nondistinctive features may then be unspecified underlyingly, and possibly even in the output of the phonology. Indeed, despite the assumptions of most researchers on Chinese tone, it is entirely possible that many syllables may leave the phonology still toneless, and have their pitch filled in by the phonetic component, as has been argued so convincingly for Japanese by Pierrehumbert and Beckman (1988). I leave this and a myriad other matters for future research.

1 This tone may be phonologically level, [L,1], and the fall may be simply phonetic.

2 A further difference lies in how many register values are available per syllable: two for Clements and Duanmu, one for everyone else. A single register value restricts the possible contours to two of a single shape, which appears to be correct. See Duanmu (1990a) for a dissenting view.

3 For Clements and Hyman (see below), H=h and L=1. They use a single feature with different geometry rather than two features.

4 Reduplication is unlikely to be involved here, since Wuyi is a Wu dialect, and such dialects typically show rightward spreading in compounds.

5 Yip (1980b,p. 170) proposes a Register spreading rule in Fuzhou, but the analysis is complex, and other explanations can be envisaged.

6 I am assuming a tonal analysis of Luganda accent. Similar examples may be readily found in other languages.

7 Duanmu does not actually use the mora. He takes all syllables to have three X-slots, one of which is for the onset, and two for the rhyme. I have translated his two rhymal X-slots into two moras.

8 Some support for this phonetic explanation comes from its apparently gradient nature. There are a few dialects in which contour tones do surface on stopped syllables, such as Pingyao and Wenzhou, and the contours tend to be gentle (54 instead of 53, 23 instead of 24, for example), as if there were an incomplete attempt to realize them on the shorter syllable. See also the F_0 tracings of one of the stopped tones in Shanghai in Zee and Maddieson (1979), which shows a slight rise. It is clear that *derived* contours can indeed surface on stopped syllables, such as the Cantonese "changed tone" formed by addition of a high tone, but Duanmu points out that these are probably lengthened, and thus would be bimoraic under his account.

9 One interesting possibility is that Register and Pitch have different TBUs. Register might have the syllable as its TBU, while Pitch has the mora. This idea has been suggested to me independently by Strang Burton and Mei-chih Laura Chan.

10 McCarthy and Prince (1986, 1990a) have argued for the existence of the Minimal Word, and suggest that it is always either bimoraic or bisyllabic. Since any syllable can be a word in Chinese, this would entail that the syllable is bimoraic, in line with Duanmu's (1990a) claim.

11 See, for example, Clements and Goldsmith (1984); Goldsmith (1982, 1987a, 1988b, 1992b); Sietsema (1989); Hyman (1978b, 1981, 1982a, 1982b); odden (1982a, 1985).

Cite this article

YIP, MOIRA. "Tone in East Asian Languages." *The Handbook of Phonological Theory*. Glodsmith, John A. Blackwell Publishing, 1996. Blackwell Reference Online. 31 December 2007 <http://www.blackwellreference.com/subscriber/tocnode?id=g9780631201267_chunk_g978063120126715>

Bibliographic Details

The Handbook of Phonological Theory

Edited by: John A. Glodsmith
eISBN: 9780631201267
Print publication date: 1996