Relational Correspondence in Tone Sandhi

by

Feng-fan Hsieh

B.A., National Chengchi University, 1996
M.A., National Tsing Hua University, 1999
M.Phil., Universiteit Leiden, 2002

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Linguistics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2007

Author ....
Department of Linguistics and Philosophy
June 29, 2007

Certified by ....
Michael Kenstowicz
Professor of Linguistics
Thesis Supervisor

Certified by ....
Donca Steriade
Professor of Linguistics
Thesis Supervisor

Accepted by ....
Irene Heim
Professor of Linguistics
Chair of the Linguistics Program

Relational Correspondence in Tone Sandhi

by

Feng-fan Hsieh

Submitted to the Department of Linguistics and Philosophy
on June 29, 2007 in Partial Fulfillment of the
Requirement for the Degree of Doctor of Philosophy
in Linguistics

ABSTRACT

This dissertation proposes that the constraint component of OT grammars should be expanded to include a family of faithfulness constraints that evaluate input-output/output-output mappings for the preservation of gross F_0 contours (rising, falling, level) across two or more segments. Following Steriade (2006), I refer to constraints in this family as Relational Correspondence constraints. The central tenet of Relational Correspondence is that phonological processes are shaped by pressure to maintain perceptual similarity between correspondent relations between successive elements, or syntagmatic contrast preservation in the auditory domain F_0, as opposed to paradigmatic contrast preservation according to which the well-formedness of an entity is evaluated with reference to the set of entities it contrasts with.

Two types of Relational Correspondence are distinguished in this work: Contour and Slope Correspondence. Contour Correspondence, formulated as RELCORR constraints, assesses correspondence of the phonological height (F_0 scaling) relation between successive tones. Four height relations are proposed for the tonal contour: "greater than" ($x>y$), "less than" ($x<y$), "equal to" ($x=y$), and "non-equal to" ($x\neq y$). Preservation of the four scaling relations is contextualized with respect to different degrees of cohesiveness: nucleus-internal, word-internal and across words. Slope Correspondence, formulated as MATCH-SLOPE constraints, requires preservation of the steepness of the F_0 contour across successive tones.

Relational correspondence provides a unifying account for a number of seemingly unrelated tone sandhi phenomena in genetically diverse languages, while explaining empirical facts that cannot be adequately expressed within the standard Correspondence Theory of faithfulness plus markedness constraints.

Thesis supervisor: Michael Kenstowicz
Title: Professor of Linguisitcs

Thesis supervisor: Donca Steriade
Title: Professor of Linguistics

TABLE OF CONTENTS

LIST OF FIGURES

Acknowledgements

I owe a profound intellectual debt to committee members: Edward Flemming, Michael Kenstowicz, and Donca Steriade. This dissertation has benefited from their guidance and insightful comments. My deepest gratitude goes to my advisors Michael Kenstowicz and Donca Steriade. I feel honored to have them as my teachers. I am particularly indebted to Donca for her generosity of allowing me to pursue the project reported here.

I would also like to thank C.-T. James Huang, Rint Sybesma, Yen-hwei Lin, Adam Albright, Shigeru Miyagawa, Shengli Feng, Suzanne Flynn, Cheryl Zoll, Jeroen van der Weijer, Douglas Pulleyblank, Moira Yip, Jaye Padgett, Curt Rice, Michel DeGraff, Duanmu San, Hsin-I Hsieh, Haruo Kubozono, Kuang Mei, Chinfa Lien, Uijin Ang, Wei-tien Dylan Tsai, Su-I Chen, Hui-chuan Hsu, Yueh-chin Chang, Wen-yu Chiang, Hsu Samuel Wang, Li-yi Huang, Liang-ji Chen for their help in these years.

My thanks also go to my classmates and friends, Sarah Hulsey, Martina Gracanin-Yuksek, Tamina Stephenson, Seth Cable, Marta Abrusan, Pranav Anand, Valentine Hacquard, Andrés Pablo Salanova, Shoichi Takahashi, Ken Hiraiwa, Sachiko Kato, Hee-sun Kim, Zhiqiang Li, Shinichiro Ishihara, Hironobu Kasai, Masakazu Kuno, and Conor Quinn, Yu-ren Peng, Han-chun Huang, Hsiao-feng Cheng, Melody Ya-yin Chang, Ryuko Taniguchi, Charles Hsu, Chi-chi Liao, Yu-chi Shen. Daphne, I hope that you will be proud of me.

This dissertation is dedicated to my parents, Ko-chao Hsieh and Chin-yuan Yang.

Finally, I hope that this piece of work shows that tone sandhi is not merely the last resort for a poor syntax/semantics graduate student who needs a topic for his or her phonology generals paper.

Chapter 1 Introduction

1.1 Central Thesis

This dissertation proposes that the constraint component of OT grammars should be expanded into a family of faithfulness constraints that evaluate input-output/output-output mappings for the preservation of gross F_0 contours (rising, falling, level) across two or more segments. Following Steriade (2006), I will refer to constraints in this family as *Relational Correspondence* constraints. The core claim of relational correspondence is stated in (1).

(1) In auditory domain *D*, a contour across two successive elements in the input should be preserved in the output.

The phenomenon of contour preservation cannot be adequately expressed wthin the standard Correspondence Theory of faithfulness (McCarthy and Prince 1993, 1994, 1995) in Optimality Theory (Prince and Smolensky 1993, 2004). This is because correspondence is conventionally defined as a relation $\Re$ between individual segments in the input string S_1 and individual segments in the output string S_2; it can be thought of as coindexation of related elements. Given the general formulation in (2), it is important to note that McCarthy and Prince's version of correspondence applies to *individual elements* of corresponding strings.[1] This type of correspondence can be dubbed "*unit correspondence*" or "*element-based Correspondence*."[2]

---

[1] Nevertheless, among the Unit correspondence constraints, LINEARITY "No metathesis" and I-/O-CONTIGUITY "No skipping/No intrusion" are exceptional in that faithfulness violations of this sort are evaluated in terms of precedence relation *between two adjacent elements* of corresponding strings (see also Heinz 2005). Therefore, we can say that the 'standard' theory of correspondence already anticipated the line of the inquiry we undertake here.

[2] See Wolf and McCarthy (2007) for a formulation of *string-based* correspondence. Note also that the Containment Theory of faithfulness also hinges on individual elements (Prince and Smolensky 1991, 1993, van Oostendorp 2005):

(2) "Element-based" correspondence (McCarthy and Prince 1995: 262)

> Given two strings S_1 and S_2, **correspondence** is a relation $\Re$ from the elements of S_1 to those of S_1. Element $\alpha \in S_1$ and element $\beta \in S_2$ are referred to as **correspondents** of one another when $\alpha \Re \beta$ (boldface in original).

This dissertation develops a theory in which preservation of identity of the contour across successive elements is required by the phonological grammar, analogous to the way that the grammar demands identity of individual elements in input-output/output-output/base-reduplicant pairs. Therefore, relational correspondence is an extension of the theory of Correspondence.

As mentioned above, the relation between successive elements is not subject to unit correspondence. That is to say, contour preservation is unexpected given unit correspondence. To see why, let us first review Steriade's (2006) core argument for the existence of contour preservation in phonological computations. The evidence comes from the robustly attested asymmetry of the different splitting potentials of TR vs. sT consonant clusters, as stated in the following implicational hierarchy, where s=fricative, T=obstruent, R=sonorant, and V=epenthetic vowel (see Broselow 1992, Fleischhacker 2001, 2005, Zuraw 2005, 2007, *inter alia*).

(3) If sT → sV(T), then TR → TV(R), but *not* vice versa,

The generalization states that cross-linguistically speaking, the sibilant-obstruent clusters (sT) are invariably less separable than the obstruent-sonorant clusters (TR). Citing data from reduplication, vowel epenthesis, infixation, loanword adaptation, and alliteration, Steriade argues that the reason sT is less splittable than TR is rooted in similarity and must be formalized as a correspondence effect. By way of example, the present asymmetry is systematically reflected in

---

"Every element of the phonological input representation is contained in the output." Note also that this definition is basically identical to Chomsky and Halle's (1968) Invariance Condition. See also Kiparsky's (1973) Alternation Condition.

reduplication of complex onsets (Steriade 1988a, Fleischhacker 2005).

(4) The TR vs. sT contrast in reduplication of complex onsets

	TR: C_1 reduplicates alone			sT: C_1 does not reduplicate without C_2			
Gothic	*gret-*	*ge-groːt*	'cry'	total reduplication:	*stald*	*ste-stald*	'have'
Greek	*grep*^{h}-	*ge-grap*^{h}	'write'	no reduplication:	*stel-*	*e-stal*	'send'
Sanskrit	*gras-*	*gaː-gras*	'eat'	C_2-reduplication:	*stu-*	*tau-stu*	'praise'

Other things being equal, C_1 in base string TR can appear in the reduplicant; however, this is not the case for sT clusters. As we can see in (4), sT clusters exhibit a greater range of outcomes. In view of unit correspondence, this contrast is puzzling in that all C_1C_2 clusters are seemingly indistinguishable from one another. More specifically, the mapping of Input C_1 to Output C_2 has no bearing on the mapping of Input C_2 to Output C_2, and vice versa.

On the other hand, markedness fails to account for the non-application of Ancient Greek reduplication: the actual output form *e-stel* should be always harmonically bound by _*se-stel_ or _*te-stel_ because, in terms of markedness, there is no *a priori* reason for both C_1 and C_2 in sT to be prevented in the reduplicant, since the C_1 of TR appears in the same environment (cf. Gouskova 2004). Moreover, markedness plays no role in alliteration (or rhyming), but such cases continue to display the sT vs. TR contrast (see Minkova 2003, Fleischhacker 2005 for details). In sum, Steriade suggests that the different separability of sT and TR is properly treated as a correspondence effect.

While a comprehensive review of alternative analyses is beyond scope of this section,[3]

---

[3] For example, a recent perceptual account advanced in Fleishchacker (2001, 2005) says that TR and TV(R) are perceptually more similar because, roughly speaking, the release burst in the transition from obstruent to sonorant creates a vowel-like percept (see also Kang 2003). In other words, TR and TV(R) are more similar than sT and sV(T) because the release burst in not available in the latter.

Steriade (2006) suggests that we pinpoint the difference between these two types of consonant clusters in terms of contour. Recall from (1) that contour, or any relation between successive elements, is manifested in an auditory dimension *D*. In the present case, the relevant dimension is intensity, which is the auditory correlate of sonority (see Parker 2002 and references cited therein). It can be seen from (5) that TR and sT differ in sonority contour shape: TR has a rising contour, while sT has a falling contour.

(5) The TR vs. sT contrast in sonority contour

	Input sonority contour	Output sonority contour	Contour Preservation
/TR/ → [TV]	Rising	Rising	Yes
*/sT/ → [sV]	Falling	Rising	**No**

If some process inserts an epenthetic vowel V into a CC cluster, then, as far as the *first two* segments are concerned, the rising sonority profile of TR is preserved in the output: /TR/ → [TV(R)], whereas the falling profile of sT is changed on the surface: /sT/ → [sV(T)]. Intuitively, if we propose a contour preservation constraint, e.g. PRESERVE-CONTOUR, requiring that an underlying contour be preserved on the surface, the ranking PRESERVE-CONTOUR » *COMPLEXONSET yields the desired result: only sT is inseparable, although the actual remedy to satisfy *COMPLEXONSET varies from language to language (see, for example, (4)). Conversely, if the markedness constraint *COMPLEXONSET dominates PRESERVE-CONTOUR, the prediction is that both sT and TR will undergo vowel epenthesis. More importantly, it is not possible for sT to be broken by epenthesis, whithout TR also being broken. The reason is the following. If a complex onset is not allowed, TR will always map to TV(R) without violation of PRESERVE-CONTOUR. In contrast, sT's mapping to sV(T) always incurs a violation of

PRESERVE-CONTOUR. Consequently, the unattested pattern in which sT is epenthesized and at the same time the TR cluster remains intact will *never* surface, according to the contour preservation account.

From the brief discussion of the TR vs. sT asymmetry, it seems that relational correspondence offers a more appealing analysis than unit correspondence.Following a smilar vein, I will explore a range of tone sandhi phenomena in this dissertation and show that the tonal contour must be regarded as an analytical primitive and thus that the auditory dimension F_0 exhibits properties analogous to those of sonority discussed in Steriade (2006), under the assumption that F_0 is the major cue for tonal discrimination (e.g. Fok-Chan 1974, Abramson 1978, Gandour 1978, Whalen and Xu 1992, Fu and Zeng 2000).

As the staring point of the overall project, I will first distinguish the predictions of these two types of correspondence by providing an overview of the empirical evidence for Relational correspondence in tone sandhi.

1.2 Why Contour Matters in Tone: An Overview

The evidence for Relational correspondence comes from the existence of phenomena which are difficult to analyze in terms of unit correspondence and/or markedness, but can be more straightforwardly accounted for if relational correspondence is taken into consideration. In this section, I highlight the role of relational correspondence in four seemingly unrelated phenomena: the optimal docking site of floating tones (§1.2.1), the adaptation of non-native F_0 contours (§1.2.2), the "invariance of variation" in tone mapping (§1.2.3), and bounded tone extension (§1.2.4). As we will see, the first three phenomena in §1.2.1~§1.2.3 involve contour preservation, or *contour correspondence*, while bounded tone extension hinges on preservation of the underlying steepness of a tonal contour on the surface, or *slope correspondence*. These two types of relational correspondence can be informally stated as follows.

(6) Two types of relational correspondence

a. Contour correspondence ≈ Preservation of the gross F_0 shape (rise, fall, level) across successive tones

b. Slope correspondence ≈ Preservation of the steepness of the F_0 contour between successive tones

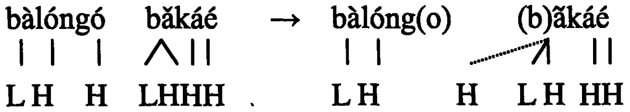
Contour correspondence assesses correspondence of the phonological height, or scaling relation between successive tones, while slope correspondence requires "steepness identity" between correspondent (contour) tones. Conceptually, contour and slope correspondence are both constructed on the key assumption that the *relations* between successive elements stand in correspondence. So, from now on, I will use the umbrella term "relational correspondence" for general purposes and specify the type of relational correspondence only when necessary.

With these discussions in mind, let us turn to an overview of the relevant phenomena in the remainder of this section.

1.2.1 The Optimal Docking Site of Floating Tones

I show in this subsection that the docking site of a floating tone is usually selected to match more closely the input contour. We will see that this observation is better understood as a contour preservation effect, hence as evidence in favor of relational correspondence.

It is well-known that one of the classic puzzles of tonology in a pre-autosegmental framework is *tone stability*, or *tone preservation*. That is, in many languages if a vowel is deleted, or undergoes glide formation, typically the tones which were associated with the deleted vowel are realized on the surface instead of also being deleted. Goldsmith (1976), drawing on Lovins (1971), illustrates tone stability with the now-classic example of Lomongo. As we can see in (7), a contraction process deletes initial consonants and vowels but the tones linked to the elided vowel are not always deleted as well. The high tone that is set afloat re-associates to the *following* vowel, resulting in a concave tone HLH. Goldsmith's insight is that tonal and segmental features belong to separate tiers so that a free tone is able to survive on the surface, even though its underlying segmental host is deleted.

(7)


'his book'

In addition to this autosegmental property, an analysis of tone preservation must also account for how the docking site of a floating tone is determined because there seems to be no *a priori* reason why the floating tone cannot dock onto the *preceding* vowel in (7). In his discussion of this example, Odden (1995: 446) notes that "[t]he H...by general convention is automatically docked on the following vowel." By "general convention," I assume that Odden means that the most common tone spreading rules are perseverative (cf. Hyman and Schuh 1974, Hyman to appear).[4] In a recent review of tonal universals, Cahill (2007) also proposes that "floating tones tend to dock rightward." Given these assumptions, a floating tone is expected to preferentially re-link to the following vowel, or tone-bearing unit (TBU). The general schema of the floating tone re-association is presented below, together with another example from Etsako, (also known as Yekhee) (Elimelech 1978). In this language, the first of two successive vowels in hiatus

---

[4] An alternative was proposed in Haraguchi (1977): a free tone normally associates with the first tone-bearing unit to its left, or, failing to do so, to its right. This convention predicts the non-existence of rightward migration of a floating tone; this prediction fails in some of the attested cases, e.g. the actual surface tone pattern in (7). See Clements and Ford (1979) for further discussion.

deletes. But in accord with the generalization mentioned above, the deleted vowel's tone is realized on the following vowel, yielding a contour tone.

(8) Rightward migration of the floating tone

b. òké òkpá	→	òkôpá	'one ram'
ówà ówà	→	ówǒwà	'every house'

It seems safe to claim that aside from other factors such as accent or syllable quantity,[5] underlying unassociated tones by default spread or dock (on)to a TBU perseveratively. So perseverativity may explain why the floating tone T_j in (8)a, which is set free due to vowel deletion, re-links to the following vowel, as evidenced in the Etsako examples in (8)b. Nevertheless, re-linking dose not always involve the following vowel. Consider a case of vowel deletion in which the docking site of a floating tone is never the following vowel (Clements and Ford 1979). In the Ewe examples below, the noun prefix *e*- deletes when following a vowel. As we can see, the tone of the deleted noun prefix *e*- re-associates to the *preceding* vowel.

(9) Ewe (Clements and Ford 1979)

ètú	'gun'	mēkpɔ̂ tú	'I saw a gun.'	
àtí	'tree'	mēkpɔ́ àtí	'I saw a tree.'	(No vowel deletion)

---

[5] For example, Steriade (1988b) notes that a number of Ancient Greek morphemes trigger deaccentuation of the stem they attach to, e.g. /a-seb-ɛ́s-ya/ → [asɛ́beia] 'impiety,' (where the high tone marking accent is represented with an acute accent). As we can see, the high tone is displaced to the preceding vowel on the surface. The present case should be regarded as tone-accent interaction.

From these examples, Clements and Ford (1979) extract a generalization, proposing the following:

(10) Stranded Tone Principle[6]

> "[T]ones which are "set afloat" due to the deletion of vowels (or, as in the case of glide formation, due to the loss of their syllabicity) reassociate to the tone-bearing unit that *conditioned* the deletion (or loss of syllabicity)." (Clements and Ford 1979: 207. Italics in original.)

In other words, the docking site of a floating tone is determined by reference to the specific processes conditioning the loss of the segmental host to which it was associated in the underlying representation. Given that the Stranded Tone Principle achieves descriptive adequacy, to the best of my knowledge, we must now try to explain why it should hold. The examples from Etsako and Ewe show that a stranded tone does not uniformly re-associate to the left or the right. In lieu of directional association algorithms, then, we might appeal to markedness, e.g. the duration-based accounts of Gordon (2001) and Zhang (2002b), since tone preservation normally results in contour tone formation. In brief, the gist of their proposals is that durationally longer syllables better accommodate contour tones. Thus, contour tones are preferentially licensed in some privileged position, for instance, phrase-final position. This is because phrase-final lengthening is a well-attested cross-linguistic phenomenon. Under this markedness view, however, the prediction is that a floating tone should dock onto the phrase-final vowel, if vowel deletion or glide formation takes place in the penult. In other words, a phrase-final lengthening account precludes the *leftward* re-association of a floating tone. As we have seen, the Etsako data in (8) and the first Ewe example in (9) do not conform to this claim. Moreover, the vowel in the deletion site is *not* lengthened, i.e. Etsako and Ewe do not have a vowel length contrast. Even if

---

[6] Notice that this term is after Kenstowicz's (1994) discussion on Margi (Hoffmann 1963).

the contour-toned vowels in (8) and (9) are provided with longer phonetic duration (presumably due to the implementation of contour tones), there does not appear to be any reason why the phrase-final vowel *cannot* do so as well. Notice further that Ewe is a Kwa language, so the well-attested penultimate lengthening of the Bantu languages does not apply. All in all, it seems fair to say that the Stranded Tone Principle is descriptively on the right track, but it is difficult to find a straightforward motivation for this generalization.

I claim that a floating tone re-links to the TBU that conditioned the vowel deletion/gliding in an effort to preserve the underlying F_0 contour. As a matter of fact, the present proposal has already been advanced in Clements and Ford (1979: 207, fn. 18), "[a] tone that has been "set afloat" reassociates to the *nearest* neighboring vowel (that is, one not separated from the deleted vowel by a consonant), regardless of direction (italics in original)."[7] Under this formulation, the competing mappings are graphically presented as follows.

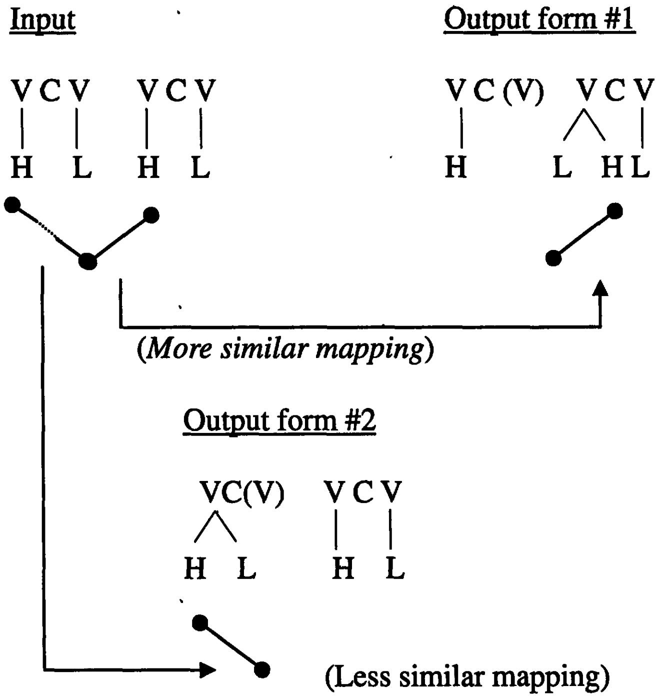
---

[7] Clements and Ford reject this formulation by remarking that it is difficult to analyze the case of floating tone reassociation due to vowel fusion (e.g. the derivation of the Kikuyu example *mwana* given in their (35), (37), and (38)). Notice that they assume that the TBU is the syllable node, represented with τ. Therefore, a floating tone is equidistant to adjacent syllable nodes, as far as the plane of τ is concerned.

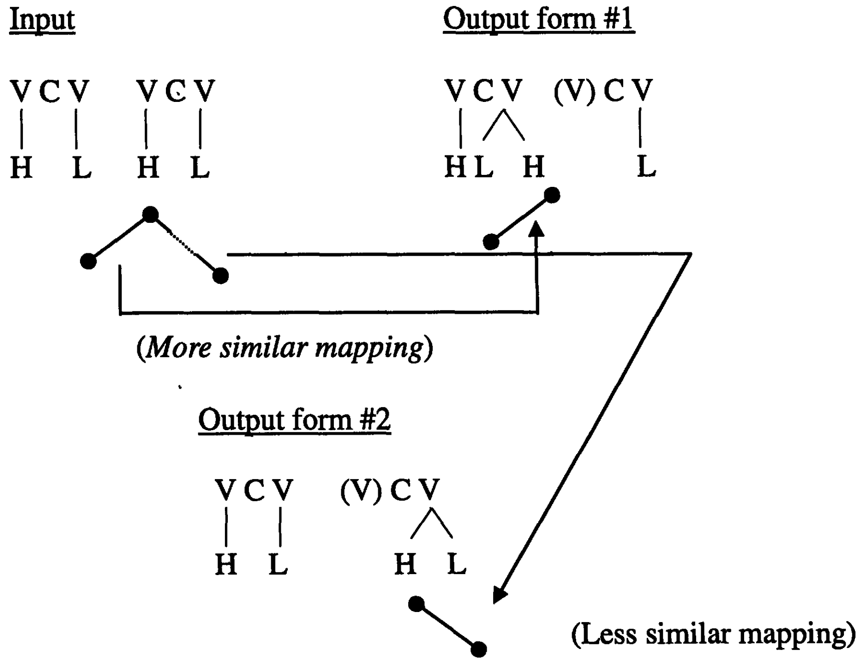
In this work, I follow Gordon (2001) and Zhang (2002b), assuming that the sonorous rime of a syllable is the carrier of tone. Under this assumption, as we will see shortly, whether the vowels in a sequence (and the tones to which they associate with) is separated by a consonant does make a difference in the selection of the actual docking site of a floating tone.

(11) Preferential retention of a tonal contour on immediately adjacent vowels

a. Rightward re-association of a floating tone (e.g. Etsako and Lomongo)



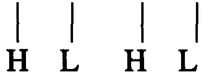
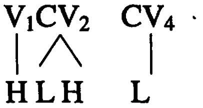
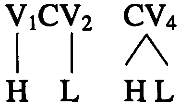
b. Leftward re-association of a floating tone (e.g. Ewe)



It is well-established that vowels, or more precisely, sonorous rimes, are better tone bearers than consonants (see chapters 2 and 3 for more discussion). Therefore, a tonal contour on a continuous vocalic span (represented with the solid line in (11)) is perceptually more salient than one that is interrupted by a consonant, especially an obstruent (represented with the dotted line in (11)). It can be seen from the above diagrams that if an underlying contour across two immediately adjacent vowels maps to an output contour on a single vowel (i.e. output forms #1), the input and output contours are more similar than the mappings in which an underlying contour interrupted by a consonant maps to an output contour on a vowel (i.e. output forms #2). It is reasonable to assume that a higher similarity value is assigned to output forms #1 because both the input and output contours are contained within a vocalic span. Therefore, the preference for output forms #1 over output forms #2 is well-grounded given that similarity is an active factor in deciding output forms. In other words, if tones are left free, they re-link to the nearest neighboring vowel that is *not* separated by a consonant, because selecting docking sites of this sort best matches the input contour across two immediately adjacent vowels.

In contrast, unit correspondence is unable to handle the Stranded Tone Principle because both rightward and leftward re-association of a floating tone satisfies every correspondence constraint that is based on individual elements, as well as the requirement of preservation of precedence structures. The point is illustrated in the tableau in (12). The following faithfulness constraints concerning precedence relation are under consideration: LINEARITY "No metathesis," I(NPUT)-CONTIGUITY "No skipping," and O(UTPUT)-CONTIGUITY "No intrusion." Suppose that the initial vowel of the second word, V_3, is deleted. It should be obvious that whether the H that is set free re-links to the left or to the right the precedence structure on the tonal tier is preserved.

(12) Indeterminacy of LINEARITY and CONTIGUITY

	LINEARITY	I-CONTIGUITY	O-CONTIGUITY
☞ a. 			
☞ b. 			

In summary, we have briefly discussed an understudied aspect of tone stability, a phenomenon classically used to support autosegmental representation: the optimal docking site of a floating tone due to the loss of the segmental host. We have seen that the existing re-association algorithms fail to generalize across all of the attested patterns. The relevant unit correspondence and markedness constraints do not fare better in this regard. On the one hand, a free tone does not by default re-link to a TBU in a specific direction. So directionality cannot be fixed with respect to floating tone re-association. On the other hand, when vowel deletion or glide formation takes place, unit correspondence is unable to select the actual surface tone patterns. I have sketched the idea that the indeterminacy of the optimal docking site, as well as the issue of direction of re-association, could be straightforwardly accounted for in terms of contour correspondence. More discussion will be provided in chapters 2 and 3.

1.2.2 Contour Preservation in Loanword Adaptation

This subsection touches on the effect of contour preservation in another seemingly unrelated domain: loanword adaptation. The core issue of loanword phonology is how non-native phonemes are modified so as to conform to the native phonotactic constraints. When there are apparent violations of native phonotactics in actual loan forms, we must seek to find the motivation behind these "exceptions." The "vowel-doubling" phenomenon in English-to-Yoruba loanwords is a good fit for the present purpose. The data are drawn from Kenstowicz (2006), which was in turn based on Ojo (1977). The following background is needed to make the case. It is well-established that Yoruba has a three-way tonal contrast (H (e.g. *rá* 'disappear'), M (e.g. *rā* 'rub'), and L (e.g. *rà* 'buy')) and features a strict CV syllable template. Abstracting away from the complications discussed in Kenstowicz (2006), we confine our attention to the adaptation patterns of the (monosyllabic) oxytone and the paroxytone loanwords from English. Kenstowicz states the core generalizations according to which the stressed syllable in English is adapted with H, while final syllable of the English source is adapted with L. The paroxytones in (13)a conform to these generalizations. In the case of the (monosyllabic) oxytones in (13)b, we see that the vowel is 'doubled.' This is the so-called vowel-doubling phenomenon just mentioned above.

(13) English loanwords into Yoruba: The paroxytones and the oxytones

a. Paroxytone		b. Oxytone	
English	Yoruba	English	Yoruba
ˈpaper	pépà	ˈbag	bá.à.gù
ˈbody	bọ́dì	ˈbat	bá.à.tì
ˈdollar	dọ́là	ˈgum	gọ́.ọ̀.mù
ˈbarber	bábà	ˈsick	sí.ì.kì

Since Yoruba lacks a vowel length contrast, at first sight it is plausible that the vowel is doubled in order to accommodate the falling H*L% contour associated with final stress in English, as noted in Kenstowicz (2006). However, in their study of the role of prosodic minimality in Yoruba vowel elision, Ola and Pulleyblank (2002) argue that minimality cannot be achieved by vowel epenthesis. More precisely, a subminimal verb such as *sè* will not be augmented by vowel insertion (e.g. *sèè* or *ìsè*) to satisfy FTBIN, the binary minimality constraint. Consequently, they conclude that the anti-insertion constraint DEP-V outranks FTBIN in the native grammar.

Turning back to the loanword data, we see two instances of vowel epenthesis. First, the English coda is adapted as CV in loan forms. This is due to the fact that the CV syllable template must be satisfied in Yoruba.[8] However a second fact, the doubled vowel in b**á.à**.gù (< bag), seems inexplicable because vowel doubling incurs a violation of the active DEP-V, and, more importantly, vowel doubling is *not* motivated by any other factors, as far as I can tell. We then want to ask why a vowel can be doubled in loan adaptation. In other words, why can't contour tones occur on a single vowel? Taking *báàtì* (< bat) for example, it is important to note that the "non-doubled" form *bâtì* will not surface, since it is well-established that surface falling/rising contours in Yoruba normally result from perseverative spreading of the preceding tone, e.g. /H.L/ → $[H.\widehat{HL}]$: /pépà/ → [pépâ] (< paper). Thus, the tone pattern in *bâtì*, $\widehat{HL}$.L, is not attested on the surface. Given that a falling (or rising) contour is not permitted to occur on a word-initial vowel, vowel doubling is the most plausible strategy to reflect the falling H*L% F_0 contour in the English source. So we are led to the following conclusion: if we subscribe to the widely accepted view according to which loanword adaptation is not independent of the native grammar, we have to find a way to overcome the DEP-V violation for the present case. As mentioned earlier, Kenstowicz (2006) notes that vowel doubling is motivated in order to faithfully realize the falling contour of the English source. This interpretation is adopted here. Therefore, it is likely that vowel doubling is driven by contour preservation, even at cost of violation of DEP-V. So

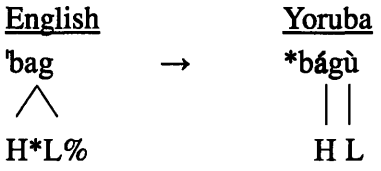
---

[8] In this regard, Ola and Pulleyblank (2002: 121, fn. 21) also note that "violation of DEP is permitted in loan adaptation."

now the question is, how can contour preservation be produced with the known OT constraints?

With respect to unit correspondence, it is conceivable that constraint conjunction (Smolensky 1993, et seq. and many others) might replicate the effect of contour preservation. Assuming that the peak H* and the final L% are perceived as H and L, respectively, we may appeal to the following conjoined constraints: MAX-(H)&MAX-(L) or IDENT-(H)&IDENT-(L). Let us first consider (14), in which the MAX constraints are employed. As we can see, MAX-(H)&MAX-(L) does not penalize the potential output form in which H and L are realized on separate syllables.

(14) MAX-(H)&MAX-(L) is satisfied as long as H and L survive.



IDENT-(H)&IDENT-(L) does not fare better, either. Recall that Yoruba lacks long vowels and hence the syllable is the TBU. It then follows that the doubled vowels in (13) are heterosyllabic. Given that IDENT-(TONE) is defined to say that the tonal specifications of TBU *x* and its correspondent *x'* should be identical, the contour preservation effect cannot be achieved: what is protected by IDENT-(H)&IDENT-(L) is HL on a *single* vowel, as shown below. This is because IDENT-(TONE) assesses tonal identity of corresponding TBUs, but the doubled vowels must be treated as two TBUs (i.e. the actual output below), as mentioned earlier. Again, the actual output forms are not produced by conjoining IDENT constraints.

(15) IDENT-(H)&IDENT-(L) forces HL on a *single* vowel

English		Yoruba	Actual output
'bag	→	*b â g ù	bá.à.gù
∧		∧ |	| | |
H*L%		H LL	HL L

From the above discussion, it is reasonable to say that unit correspondence seems a poor fit for the phenomenon in question. By contrast, the contour correspondence analysis in the preceding section can be extended to the English-to-Yoruba loanword data without any problem. Since the English H*L% contour is realized on a continuous vocalic portion in *bag*, the corresponding HL tone sequences should also be contained on a continuous vocalic portion. It is obvious that vowel doubling is the optimal strategy to meet this requirement, albeit at the expense of violating DEP-V. Therefore, the desired results are achieved if we rank the contour preservation constraint (tentatively called "PRESERVE-CONTOUR") over DEP-V.[9]

This completes our discussion on the contour preservation effect in loanword adaptation. To recapitulate, the puzzle is that when confronting a non-native falling contour, speakers tend to preserve the falling contour by vowel doubling, which is surprisingly not permitted in the native grammar. I have shown that it is difficult to analyze this phenomenon in terms of unit correspondence. Once contour preservation is taken into account, however, we see that the English loanword data have this key affinity with the Stranded Tone Principle: preservation of the tonal contour within a vocalic span.

---

[9] One may wonder if under the assumption of the Richness of the Base, the present analysis wrongly predicts that an underlying H͡L associated to a vowel will surface as a falling-toned vowel. Notice, however, that Yoruba does *not* have underlying contour tones. There are only three level tones in the inventory, H, M, and L, so H͡L will not be included in the tone inventory (presumably due to insufficient distinctiveness). In other words, H͡L will never be present in the input. Hence, contour preservation is vacuously satisfied. See Flemming (2006) for the formulation of this model.

1.2.3 Invariance of Variation in Tone Mapping

Additional evidence lending support to contour correspondence comes from a phenomenon I will term "Invariance of Variation." For the present purpose, it should be sufficient to present the core argument with the following schematic example. Detailed discussion of a full array of data in Shanghai Chinese will be provided in §2.3.

The schematic example is as follows. Suppose that *x*, *y*, *a*, and *b* are tone-bearing units, *a* is the output correspondent of *x*, and *b* is the output correspondent of *y*.

(16) *Invariance of variation in tone mapping*

Input		Output	Remarks
x *y*		*a* *b*	Unfaithful in unit correspondence
$\vert$ $\vert$	→	$\vert$ $\vert$	*Faithful in relational correspondence*
H M		M L	
(Falling contour)		(Falling contour)	

We see in (16) that an underlying tone sequence HM surfaces as ML in the output. Let us follow McCarthy and Prince's (1995) suggestion and assume that correspondence relations may also hold between other kinds of elements, in particular, features and prosodic units (see, for example, Lamontagne and Rice 1995, Lombardi 1995, 1998, 2001, Pulleyblank 1996, Causley 1997, Walker 2000, Zhang 2002a). It then appears that the mappings in (16) incur violations of three of the major correspondence constraint families.[10]

---

[10] The other unit correspondence constraint families include: UNIFORMITY "No coalescence," INTEGRITY "No breaking," I-/O-CONTIGUITY "No skipping/No intrusion," LINEARITY "No metathesis," and I-/O-ANCHORING-(L/R) "No deletion/insertion at the left/right edge to the input/output string." It should be obvious that these constraints are all irrelevant here.

(17) Major correspondence constraint families (McCarthy and Prince 1995)

Max	"Every segment/feature in S_1 has a correspondent in S_2."
Dep	"Every segment/feature in S_2 has a correspondent in S_1."
Ident	"Correspondent segments are identical with respect to feature [F]."

Max requires every segment/feature in S_1 to have a correspondent in S_2, or "no deletion," but we see in (16) that neither H nor M in the input survives in the output. Dep, on the other hand, penalizes insertion. That is, any segment/feature in S_2 should have a correspondent in S_1. As seen in (16), the M and L on the surface do not have correspondents in the underlying representation. Finally, featural identity in the mappings H → M and M → L is also inexact. Taken together, the above discussion suggests that the mappings in (16) should not occur according to unit correspondence. Thus, as a first approximation, it is fair to say that the mappings in (16) are difficult to analyze within purely element-based correspondence constraints.[11]

Instead, (16) *is* a faithful mapping with respect to the *contours* across HM and ML. This is because the falling contour is preserved in the output even though each tone has changed, be it element-wise or specification-wise, on the surface. More precisely, both of the sequences HM and ML form falling contours, although the two falling contours differ in their starting and ending points in tone value. In other words, what remains constant in (16) is the falling profile of the contours in the input and the output representations, as the term "Invariance of Variation" suggests.

In conclusion, although our discussion of the above schematic example still needs to be justified, it is appealing to note that contour correspondence again makes very different predictions from unit correspondence: contour tone could be protected even at the expense of

---

[11] It is conceivable that HM can be treated as a H-register falling tone and ML a L-register falling tone (modulo the vexed problem of M's register specification). So the issue in question could be analyzed as change of tonal register feature specification (Yip 1980, Bao 1990, among many others): HM (H-register) changes to ML (L-register). See § 2.3.3 for discussion against this alternative approach.

violating every constraint on correspondent elements. See §2.3 for real examples and relevant discussion.

1.2.4 Bounded Tone Extension

The last case in this overview section is a phenomenon I termed 'bounded tone extension.' As mentioned at the outset, the phenomenon in question involves a different instantiation of relational correspondence, i.e. slope, or the degree of the steepness of tonal contour. I assume that slope is defined as the ratio of F_0 differences between two tones (i.e. the altitude change) to duration between two tones (i.e. the horizontal difference).

(18) $$\text{Slope} = \frac{F_0 \text{ difference}}{\text{Duration}}$$

As a reminder, we have pointed out that contour correspondence concerns preservation of the correspondent phonological height, or scaling relation, between successive tones. In the same vein, *slope correspondence* hinges on preservation of the degree of the steepness between successive tones in the input and output.

I show in this subsection that the evidence for slope correspondence comes from the other side of the coin of *tone mobility*. That is, unlike most segmental features, it is well-known that tone may move several syllables away from its lexical source. An oft-cited example concerns Chizigula (Kenstowicz and Kisseberth 1990), where high tone migrates from the verb root to the metrically strong penultimate syllable of the word, as the following data illustrate. Note that low tone is unmarked.

(19) Tonal attraction to the penult in Chizigula (Kenstowicz and Kisseberth 1990)

lómbez 'to request' → ku-lombez-ez-án-a 'to request for each other.'
| |
H   H

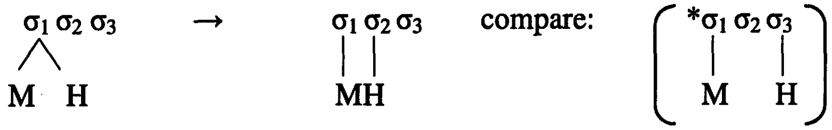
A sometimes overlooked observation is that unbounded tone displacement of this sort is generally attested in the case of *single* tone. Tone mobility is much more restricted in contour tones. As we know, contour tone is conventionally treated as a composition of at least two level tones. So the restriction can be construed as a "locality" constraint: informally speaking, the underlying tonal sequence should be as "close" as possible on the surface. For the present purpose, it suffices to consider some representative data from Shanghai Chinese since bounded tone extension is amply attested in polysyllabic tone sandhi in Wu Chinese (and in a handful of Mandarin dialects). Some background is provided as follows. The general schema of Shanghai Chinese tone sandhi is that all tones of a phonological phrase are neutralized (here I use a theory-neutral term "tone loss"), except those of the initial syllable, abstracting away from the case of the checked rising tone, to which I shall return in §2.3 and §5.5.1. The surviving tones are then redistributed over the *first two* syllables, regardless of the syllable number of the phonological phrase. The facts are illustrated by the following data in which H-register rising tone MH occurs on the initial syllable (Zee and Maddieson 1980). Note that tone loss is represented with the symbol '∅' and the upstep symbol '↑' marks a raised tone.

(20) Shanghai Chinese tone sandhi (Zee and Maddieson 1980: 46)

UR	Tone loss	SR	Example	Gloss
MH-MH	MH-∅	M-ꜜM	ɕiɔ tsɔ	'portrait'
MH-MH-MH	MH-∅-∅	M-**H**-L	ɕiɔ tsɔ ɕiã	'small photograph'
MH-MH-MH-HL	MH-∅-∅-∅	M-**H**-M-L	ɕiɔ tsɔ ɕiã tɕi	'small camera'

The pitch contour of the third and the fourth syllables can be treated as an interpolation of the H peak and the low boundary tone (L%).[12] More importantly, it appears that the realization of MH is subject to a locality requirement once tone redistribution takes place: the H peak (in boldface) of the initial underlying MH is invariably realized on the second syllable. In other words, tone redistribution in Shanghai Chinese cannot be unbounded: tone is unable to migrate from its underlying host to a syllable that is not adjacent to its lexical source in (21).

(21) Bounded tone extension



This contrasts with Chizigula where a singleton H tone is attracted to the penult even though the penultimate syllable is several syllables away from H's lexical source. What could be the driving force behind this typological disparity? Some discussion is in order. First, it is hard to see how this observation can possibly be captured with unit correspondence. MAX/IDENT-(TONE) is unable to select the optimal output because both the bounded and the unbounded (in parentheses) tone extensions satisfy or violate MAX/IDENT-(TONE) equally. Second, precedence structures are preserved from the input to the output tone sequences and vice versa. Thus, I-/O-CONTIGUITY is

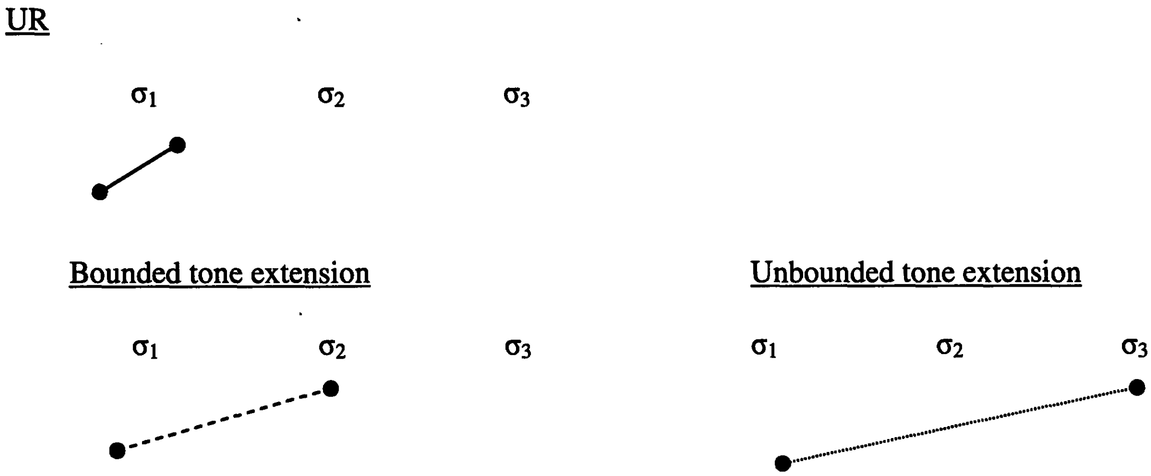
---

[12] As for the disyllables, let us simply assume for now that the H peak merges with L%, yielding Mꜜ.

also irrelevant in this regard.

When we look more closely, however, another possible analysis suggests itself. Recall from (18) that slope is defined as the following ratio: F_0 difference over duration. Given this, we see in the diagram in (22) that the steepness of the bounded tone extension (i.e. the dashed line) is more similar to the input slope (i.e. the solid line), than is the unbounded tone extension (i.e. the dotted line), whose slope value is the shallowest among the three contours. Suppose that there is a slope correspondence constraint, requiring that the input and output steepness of the corresponding contour tones should be, roughly speaking, as similar as possible. It then follows that the H peak in (20) and (21) is realized on the second syllable. In other words, bounded tone extension is motivated in order to maintain the similarity of correspondent slopes.

(22) Bounded tone extension



UR

Bounded tone extension

Unbounded tone extension

Furthermore, the unbounded tone displacement in Chizigula (and in other Bantu languages) also lends support to slope correspondence from another angle. We can now postulate that the single tone is able to migrate from its lexical source to a remote host because the requirement of slope-matching will be always satisfied as long as a level tone remains level on the surface.

In summary, our discussion suggests that it is beneficial to consider tone sandhi in terms of slope-matching. As we have seen, it is typically the case that the members of an underlying contour tone are not separated by a syllable on the surface. The point here is that bounded tone extension does not seem explicable by unit correspondence because faithfulness of *the elements on the tonal tier* will not be changed even if tone extension is unbounded (cf. (21)). In contrast, it appears that slope correspondence serves as a more straightforward motivation for bounded tone extension. More discussion will be provided in §5.5.

1.2.5 Summary of this Section

In this section, I have briefly discussed four seemingly unrelated tonal phenomena: i) the optimal docking site of a floating tone, ii) the adaptation of non-native F_0 contours, iii) the "Invariance of Variation" in tone mapping, and iv) bounded tone extension. I have shown that these phenomena are difficult to analyze with the known OT constraint families, in particular, with the standard theory of faithfulness: unit correspondence. Instead, their analyses are better elucidated by relational correspondence, or, more specifically, contour and slope correspondence, i.e. contour and slope-matching constraints.

Before we move on to the formalization of the relational correspondence constraints in tone and tone sandhi, it is necessary to spell out the representational framework of tone adopted in this dissertation, to which I turn in the following section.

1.3 Empirical Background and Representational Issues in Tone

In this work, I adopt a scalar representation of tone, following Flemming's (1995, 2002) multi-dimensional auditory space. The dimensions are multi-valued features, and I will present F_0 in terms of a five-point scale. Simply put, I basically employ Chao's (1930) tone letters to represent tones (with an additional specification, slope, to which I shall return in §1.3.2). Most current theories of tone, by contrast, assume a binary approach to tonal features (see Bao 1999, Chen 2000, Yip 2002 for comprehensive overviews of various models). Under the conventional wisdom of binarity, the tonal space is divided into two subranges, the H- and L-registers, represented with [±U(pper)], or comparable laryngeal features, while tone is specified with a binary feature [±High], or equivalent laryngeal features, represented with H/L throughout. Let us call this the "standard" feature system. Some well-acknowledged problems with a binary system of tonal representations can be understood by looking at Yip's (2002) proposal regarding the desiderata for a feature system for tone. Of present interest is the following:

(23) Desiderata for a feature system for tone (Yip 2002: 40)

i. Characterize all and only the numbers of level tone contrasts (= her (2a))
ii. Characterize contour tones, and their relationship to level tones (= her (2b)
iii. Characterize all and only the number of contour tone contrasts (= her (2c))

The first problem lies in the fact that there are languages with five contrasting level tones. See Maddieson (1978: 338) for a list of languages that have been described as having a five-level tone inventory. Notice that it has been attested that not all such tones are produced with modal phonation. For example, Zhu (2006) reports that the highest level tone, 55, in Gaoba Dong (Tai-Kadai) co-occurs with falsetto phonation only, while the other four level tones, 11, 22, 33, and 44 are produced with modal voicing (see (24) for more discussion). So the four level tones can be dealt with in the standard feature system. Nevertheless, Ziyun Miao (Hmong-mien), another language making five-way contrasts in level tones, does not have this property (Kong 1992). More crucially, with regard to Maddieson's (1978: 338) now-classic claim: "language may contrast up to five levels of tone, but no more," it is puzzling why the number of level tones should be fixed in this manner, given a binary feature system, or more generally, given a categorical feature approach. In other words, an (as-yet-unattested) six-level tone inventory is presumably equally possible provided that tones are represented with a set of these features, contrary to fact (a point also made by Myers and Tsay 2003). Similarly, another perennial difficulty of the standard model is generating the potentially infinite number of levels in a series of downsteps (see, for example, Clark 1978, Clements 1979, Hyman 1993, Snider 1990, 1999).

Secondly, a binary model is unable to handle the "too many contour tones" problem. One of the famous examples comes from San Andrés Chicahuaxtla Trique (Hollenbach 1977, see also Yip 2002). This language has four level tones and contrasts five falls and four rises on final syllables: 12, 23, 34, 35, 45 and 21, 43, 53, 54 (note that 1 is the highest and 5 the lowest). It should be obvious that a binary feature analysis is a poor fit for such a system. As a matter of fact, the "too many contour tones" problem is not uncommon in Sinitic languages as well. See, in particular, chapter 4 for the case of Shaoxing Chinese, a Wu Chinese language with four falling tones in citation, 551, 232, 31, and 221, and §5.3 for Hangzhou Chinese, a Wu Chinese language having two L-register rising tones in isolation, 13 and 23.

By way of a concrete example, let us observe the F_0 tracks of citation tones in Lishui Chinese, a Southern Wu Chinese language. The data are taken from Sheng's 2001 acoustic study. This language has five citation tones, and among them there are *two* H-register falling tones, 551 and 42. One might contend that 42 can be treated as a mid-falling tone, hence is represented as [-U, HL], contrasting with 551, whose tonal specification is presumably [+U, HL]. However, this treatment runs into problem with tone 232, which is also a mid-falling tone, whose initial dip is due to the F_0-suppressing murmured onset, a wide-spread areal feature in Wu Chinese (see §2.3.1 for more details). Notice further that murmur (or breathy voicing) in Wu Chinese is *not* a tonal feature (contra Yip 1993), as extensively discussed in Ren (1992). Simply put, it is not always the case that L-register tones co-occur only with breathy voicing. A good example again comes from tone 21 in Lishui Chinese in Figure 1-1: tone 21 is historically of the Yin register, whose synchronic reflex is modal phonation (i.e. non-murmur), but it is obvious that tone 21 is positioned in L-register, indicating that murmur does not absolutely correlate with L-register (at least in Lishui Chinese). Therefore, murmur does not tonally distinguish tone 42 from 232. In other words, the present discussion leads to an undesired conclusion for the standard theory: 42 and 232 should be both specified with [-U, HL].

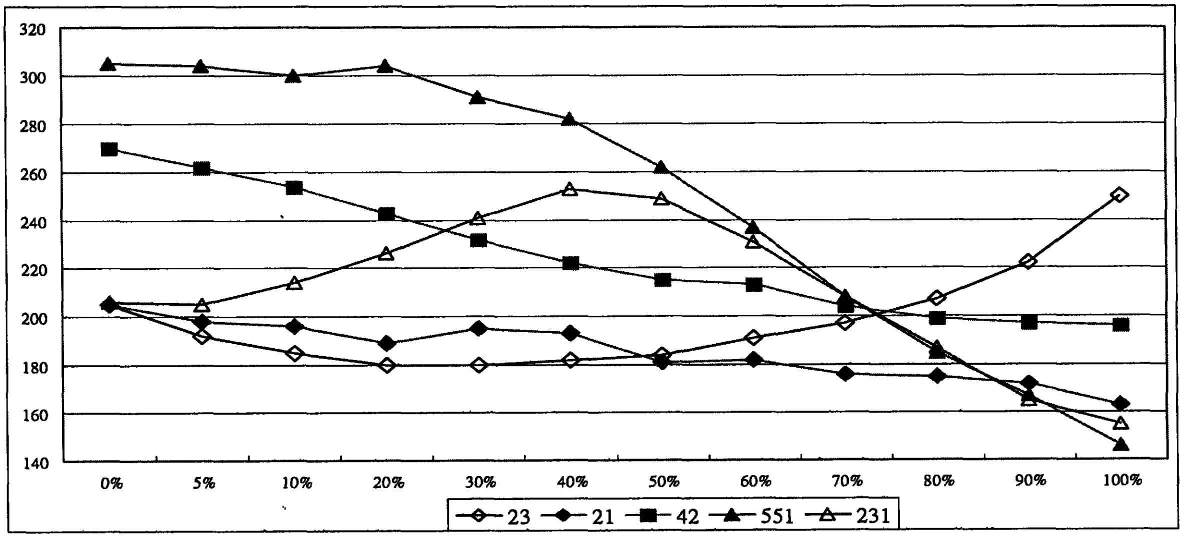


Figure 1-1 Normalized F_0 (in Hz) for long (non-checked) citation tones in Lishui Chinese

The Lishui Chinese data also bring up a third problem for a binary approach to tonal features: the "overlapping contour" problem. More precisely, an H-register falling tone is conventionally specified with [+U, HL].[13] Such a tone can be transcribed as 53 in Chao letters because a H-register low tone is realized as a mid tone on the surface. If a language has an H-register falling tone, however, this high falling tone is typically realized as 51, i.e. a contour ranging from an H-register high tone (i.e. 5) to an L-register low tone (i.e. 1). This is not a possible featural combination given that there is only one register feature per tonal node. As we have seen in Figure 1-1, Lishui Chinese has two falling tones, 551 and 42. In view of contrast preservation, we expect that 53 and 42 should be a better pair in terms of contrast for these two falling tones in Lishui Chinese. But the final portions (i.e. from 50% to 100% normalization points) of their F_0 contours largely overlap. It should be now obvious that the overlapping contour problem cannot be easily accommodated in the standard model.

To my knowledge, 51 (or high-to-low) and 53 (or high-to-mid) are not contrastive in any language. High and low falling tones are usually, if not always, phonetically realized as 51 and 31 (cf. the case of Lishui Chinese in Figure 1-1). One may contend that there could be a "universal phonetic rule" that converts the phonological form [+U, HL] into the phonetic form 51. Under this assumption, 53, or a high-to-mid fall, should be unattested on the surface. Consider, however, the following tone inventory of Hangzhou Chinese (data from Huang's 2001 acoustic study; see also §5.3 for more discussion), where a high-to-mid fall *is* indeed attested (i.e. 433). The present data indicate that [+U, HL] can be realized as 53, or a high-to-mid fall, at least in Hangzhou Chinese, suggesting that the purported "universal phonetic rule" fails to apply in this regard. That

---

[13] See Duanmu (2000) for a categorical feature-based treatment. Note that in his system, [slack] is associated with breathiness, which co-occurs only with L-register tones. Under this formulation, then, L-register tones should not occur both in the modal and the breathy phonation. Moreover, it is not clear how creaky voice can be compatible with L-register tones.

being the case, there does not appear to be a principled motivation for this particular "universal" phonetic rule. In sum, the "overlapping contour" problem also seems problematic for the standard model.

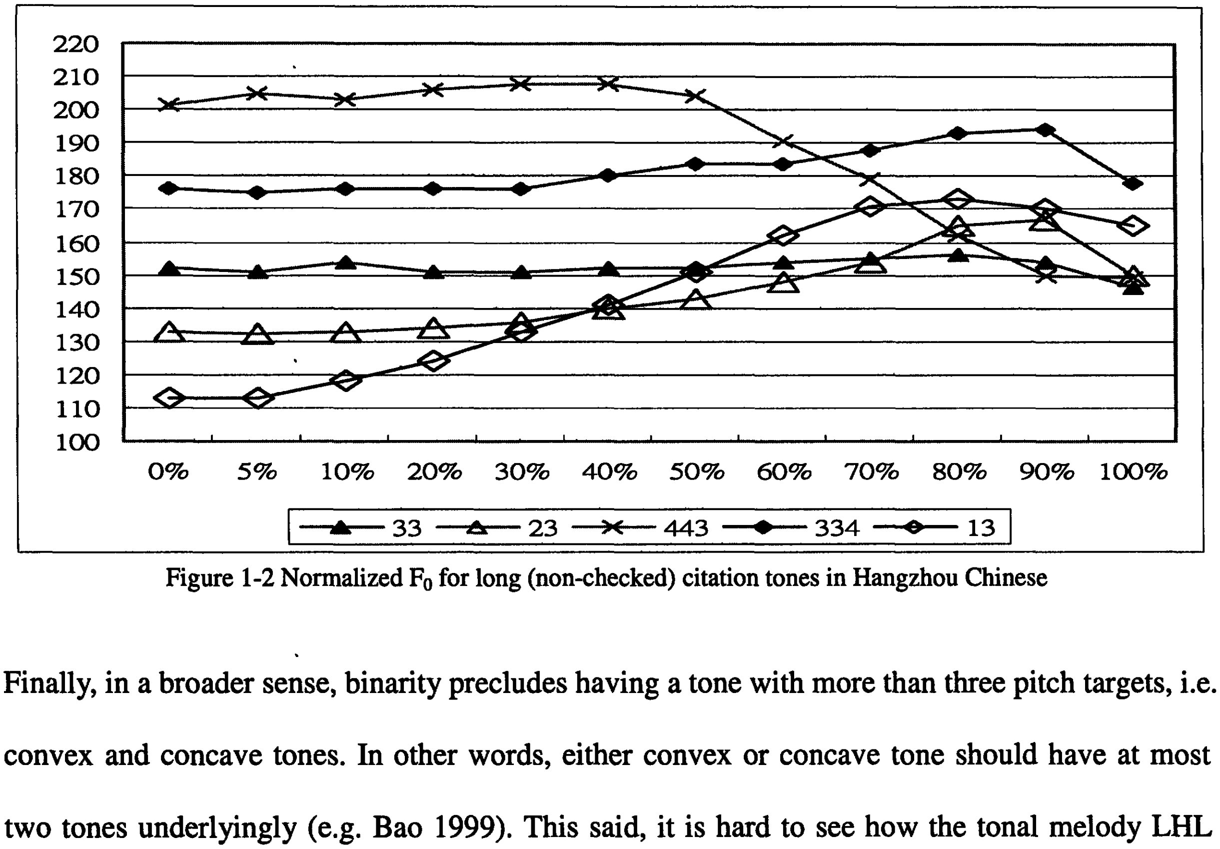


Figure 1-2 Normalized F_0 for long (non-checked) citation tones in Hangzhou Chinese

Finally, in a broader sense, binarity precludes having a tone with more than three pitch targets, i.e. convex and concave tones. In other words, either convex or concave tone should have at most two tones underlyingly (e.g. Bao 1999). This said, it is hard to see how the tonal melody LHL can be realized on a monosyllabic word in Mende (Leben 1973), for example.

Based on the above discussion, a gradient approach to tonal representation seems to fare better. On the one hand, the number of level tones is not necessarily limited to four. On the other hand, both the "too many contour tones" problem and the "overlapping contour" problem can be easily handled by a gradient approach. Gradient approaches to tonal representations are far from novel (e.g. Snider 1990, Tsay 1994, Zhang 2002b, Myers and Tsay 2003). The advantage of a gradient approach is, however, not widely acknowledged, presumably due to the "overgeneration" problem. More specifically, one may wonder, for example, why there is no attested language that has an inventory with 25 contour tones, given a five-level system. As discussed in Yip (2001), Zhang (2002b), and Myers and Tsay (2003), overgeneration of the unattested tonal contrasts can be in principle constrained by the Dispersion Theory (Flemming 2006, 1995, 2002). While developing a theory of tonal dispersion is beyond the scope of this dissertation, I would like to argue for a scalar representation of tone by addressing two sometimes overlooked issues in the relevant literature. The first issue concerns the (possible) logarithmic nature of tone production and perception. I will show in §1.3.1 that this specific property may shed light on the well-established "five-level" constraint on tonal contrasts (see also Tsay 1994 for an extra-grammatical account based on, in particular, memory constraints). Second, I will show §1.3.2 that pitch shapes *are* phonologically contrastive. This view has been anticipated in Pike (1948); he raises the possibility that it might be necessary in some languages to distinguish tonal contours not only in terms of their end points and/or direction, but also in terms of such features as rate of fall or rise, correlation between "time and distance of rise," etc. As we will see, contrasting pitch shapes cannot be easily expressed in the standard approach of tonal representation.

1.3.1 The (Possible) Logarithmic Nature of Tone Production and Perception

The goal of this section is to show that tone production and perception may be logarithmic, but not linear, at least in some languages of the world. This property is reflected in the fact that pitch intervals are not always evenly spaced, as one might otherwise expect. In other words, given a three-level tone system, it is *not* always the case that the distance between the high and the mid tone is the same as the distance between the mid and the low tone. Instead, in some languages, there is a general trend according to which pitch levels are usually wider in the higher part of the pitch range. In a five-level tone system like Gaoba Dong (Tai-Kadai) in (24), pitch intervals are obviously narrower between lower tones. Notice that according to Zhu (2006), 55 is pronounced with falsetto phonation, which may explain why 55's average F_0 value is extraordinarily high. 20 tokens were measured (5 tones with 4 monosyllables, obtained from one male speaker SL), using KAY 7030 (Shi et al. 1987).

(24) Gaoba Dong (Shi et al. 1987)

Tone value	Hz	Difference in Hz between each tone and the lowest tone	Phonation type
11	129	0	Modal
22	138	9	Modal
33	157	23	Modal
44	197	68	Modal
55	259	130	Falsetto

Similarly, we can see in Figure 1-3 that tone levels are also not evenly spaced in Copala Trique (based on two speakers, Hollenbach 1984): level tones are represented with the solid line, and contour tones with the dotted line. Again, the lower tones are closer to one another. The rising/falling onglides of the highest/lowest tone are cross-linguistically not uncommon in utterance- or word-final position (Maddieson 1978: 341). Presumably, these onglides help speakers to identify the tone correctly.

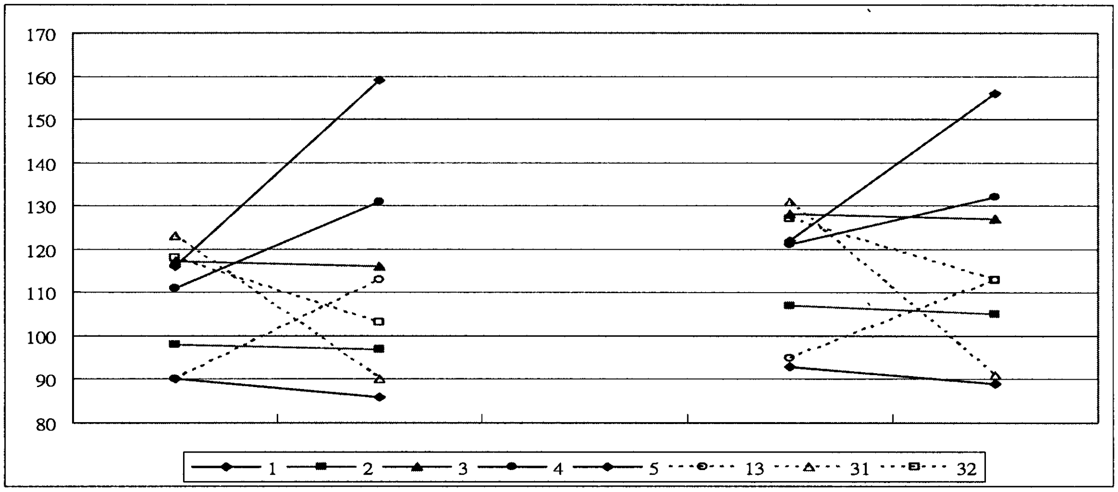


Figure 1-3 Average F_0 (in Hz) for Copala Trique citation tones (Hollenbach 1984: 73, Table 5)

Finally, the F_0 differences between the five level tones in Ziyun Miao (Hmong-mien) are presented below (Kong 1992). 18 tokens (obtained from one male speaker) were measured, using KAY 7800. As we can see, pitch intervals are not evenly spaced here either. In particular, the lower tones are again closer to one another than the higher tone are.

(25) Ziyun Miao (Kong 1992)

ΔF_0(11-22)	12 Hz
ΔF_0(22-33)	18 Hz
ΔF_0(33-44)	21 Hz
ΔF_0(44-55)	18 Hz

Let us now consider the relevant data from two four-level tone systems. The Manbila data also conform to the present generalization (Connell 2000): pitch intervals of lower tones are narrower. Mean F_0 values for the beginning, middle, and end points of the four level tones (from 3 speakers) are plotted in Figure 1-4. The rising onglides of Tone 1 and the falling onglides of Tone 4 can again be understood by Maddieson's (1978) tonal discrimination account. This is also in line with the results of Harrison's (1996) perceptual experiments, suggesting that low and mid tones are distinguished by a low tone's falling contour in Yoruba (cf. LaVelle 1974, Hombert 1976).

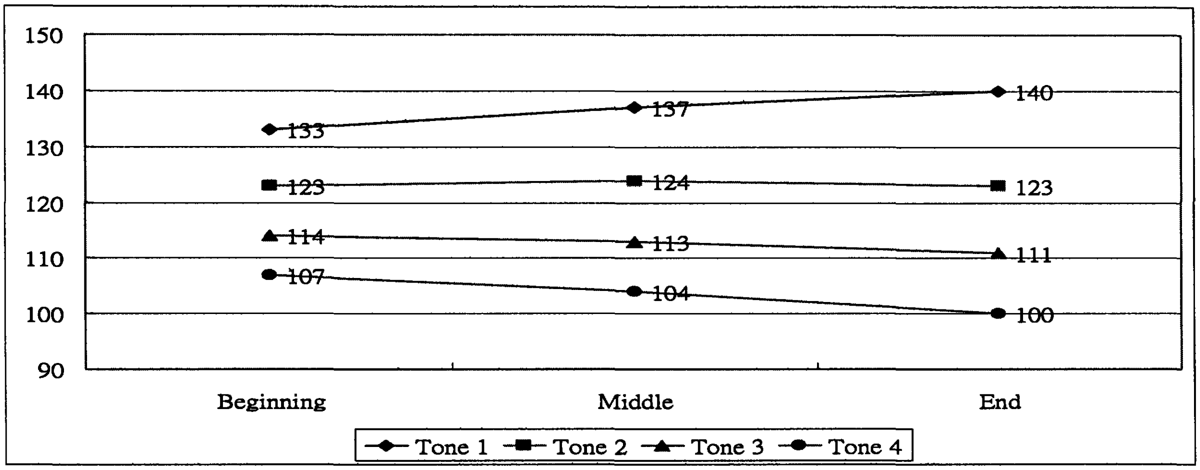


Figure 1-4 Average F_0 values (in Hz) for four level tones in Manbila

Cantonese is another tone system making contrasts of four levels. The following pictorial depiction approximates normalized F_0 contours for Cantonese tones reported in Lau (2000), based on 1,800 monosyllabic words produced in isolation by one male speaker. Long (non-checked) tones are represented with the solid line, while the dotted line represents checked tones. We see that a recurrent pattern emerges: pitch level in the higher part of the F_0 range is much wider.

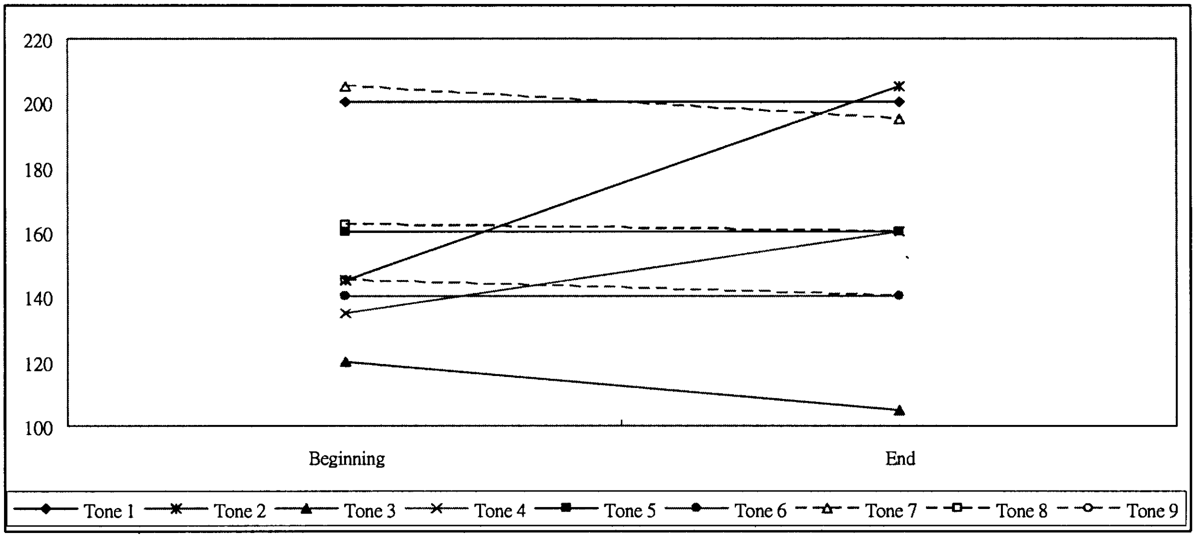


Figure 1-5 Normalized F_0 (in Hz) for citation tones in Cantonese (replicated in Lau 2000)

A similar distribution of pitch interval is attested in a three-level system such as Buli (Gur) or Tianjin Chinese. Akanlig-Pare and Kenstowicz (2002) report that high tone in Buli is significantly higher in F_0 value than mid and low tones (data from one male speaker).

(26) Buli (Akanlig-Pare and Kenstowicz 2002)

Low	100 Hz
Mid	130 Hz
High	150~200 Hz

Likewise, the three-level pitch heights in Tianjin Chinese also show that the F_0 difference between M and L is smaller than that between M and H by 20 Hz. Forty tokens (4 tones with 10 monosyllables) were measured, using Praat (Boersma and Weenink 2007). Data are recordings of one female speaker extracted from Yang et al.'s (2004) *The Phonetic Database of Tianjin Chinese.*

(27) Tianjin Chinese (personal unpublished data)

Low	300 Hz
Mid	220 Hz
High	160 Hz

It seems that tone levels are asymmetrically narrower in the L-register, at least in the languages discussed above.[14] We then should try to find a motivation for the phenomenon in question. It has been noted that the F_0 distribution is not normal as one might expect, but rather is positively skewed towards the lower part of a given F_0 range (e.g. Zhu 1999 and reference cited therein). Positive skew means that the mass of the distribution is concentrated on the left of the figure in (28), or, in other words, more data on the right tail than would be expected in a normal distribution.

(28) Positive skew

Regarding F_0 distribution, positive skew refers to the generalization according to which the lower part of the pitch range is usually, if not always, more exploited in making tonal contrasts. By way of concrete examples, normalized F_0 tracks of citation tones in two Wu Chinese languages are presented below. The first example comes from Wenzhou Chinese. Data were replotted from Sheng's (2001) acoustic study, which is based on one male speaker SCH, aged 63 in 2001. 9 tokens (3 monosyllables with 3 repetitions) were measured from each tone, using KAY CSL 4300B. Note also that checked tones (i.e. tones on glottal stop-terminating syllables) are

---

[14] Note that pitch intervals are evenly spaced in Yoruba and Thai according to data cited in Maddieson (1978).

underlined.[15]

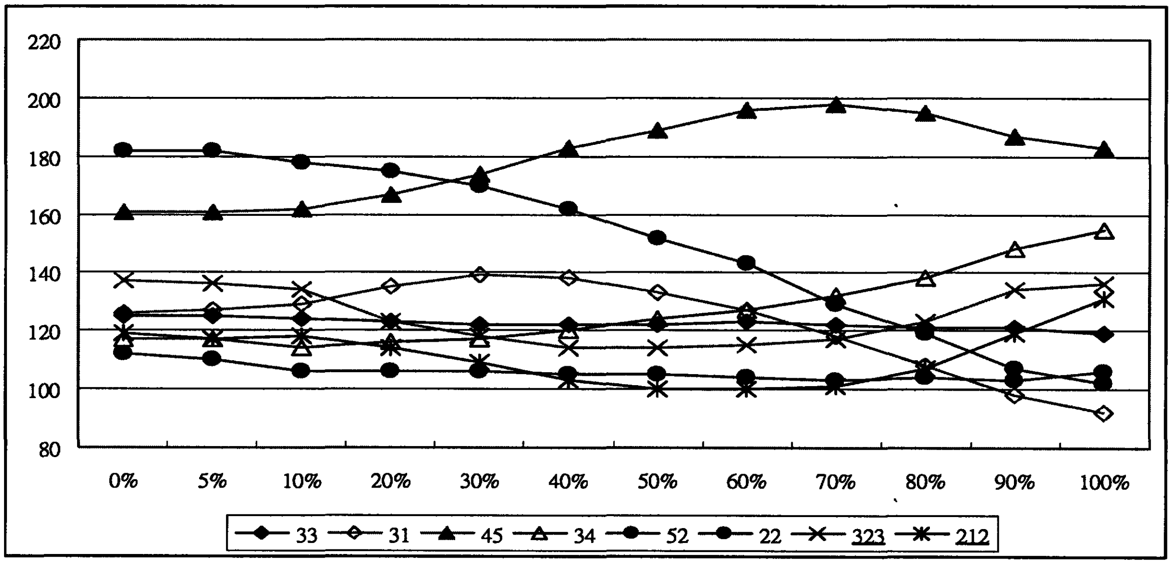


Figure 1-6 Normalized F_0 (in Hz) for Wenzhou Chinese citation tones

Wenzhou Chinese has a citation tone inventory of eight tones. It should be obvious from the above pitch tracings that tones crowd in the lower part of the F_0 ranges (roughly below 140 Hz), while less contrasting tones are positioned in the higher part. In addition to this, normalized F_0 contours of another eight-tone system, Shaoxing Chinese, are illustrated as follows (see chapter 4 for more the analysis of disyllabic tone sandhi in Shaoxing Chinese). Data are taken from Ping's (2001a) experimental report, based one male speaker TH, aged 32 in 2001. 15 tokens (5 monosyllables with 3 repetitions) were measured from each tone, using KAY CSL 4300B. Similarly, the Shaoxing Chinese data also conforms to the property of F_0 distribution observed above: positive skew.

---

[15] Incidentally, Wenzhou Chinese might be the only Sinitic language in which checked tones are longer than non-checked tones.

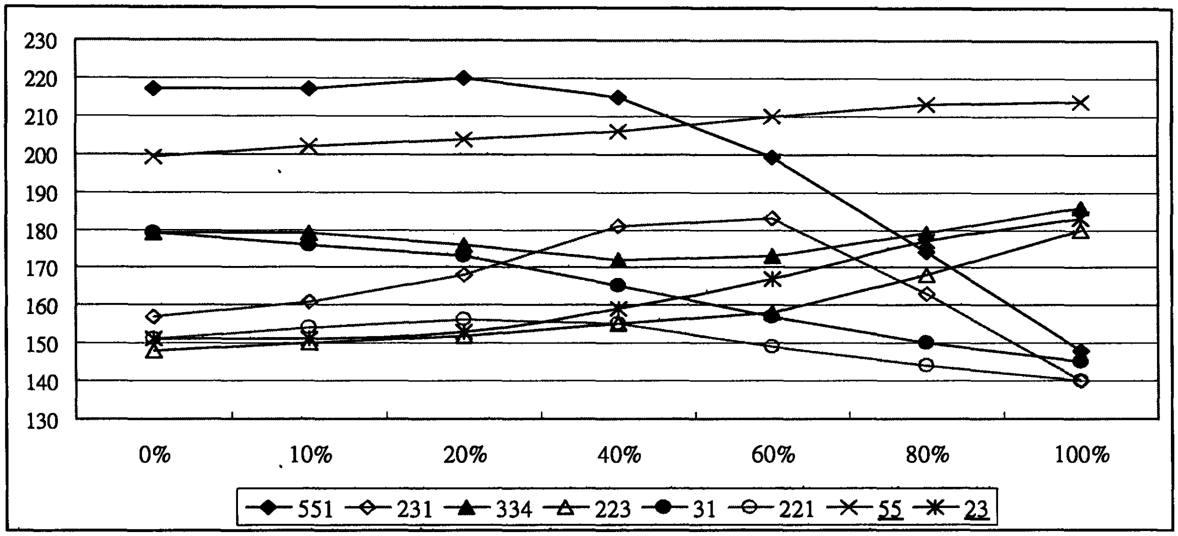


Figure 1-7 Normalized F_0 (in Hz) for Shaoxing Chinese citation tones

The positive skew is attested both in atonal and tonal languages.[16] This probably reflects the logarithmic nature of the F_0 change according to which "the mechanical motion of an element in the laryngeal mechanism which from the point of view of pitch control, can be approximated by a second-order linear system [i.e. decaying oscillations]" (Fujisaki 1983: 52). More precisely, "a smaller F_0 interval in the lower part [of the pitch range] has the same distance on the logarithmic scale as a larger F_0 interval in the higher part" (Zhu 1999: 54).

On the perception side, it is well-known that the mel scale is logarithmic over 1 kHz. But it is controversial as to whether the mel scale is logarithmic or linear below 1 kHz. Zhu (1999: 54-55) is of the opinion that the mel scale is also logarithmic below 500 Hz (albeit approximately linear). While I am unable to offer any fresh insight or new contribution to this issue, I would like to stress the following point. Turning back to Wenzhou and Shaoxing Chinese, if the perception of tone does not work on a logarithmic basis, it is unclear why the lower part of the pitch range should be more exploited. In addition to this, as we have seen above, similar

---

[16] According to Zhu (1999: 54), atonal languages include: Polish (Jassem 1971) and English (Menn and Boyce 1982: 379); tonal languages include: Shanghai Chinese (Zhu 1999), Vietnamese (Earle 1975: 153) and Pakphanang Thai (Rose 1994).

distribution of pitch levels has been attested in genetically unrelated languages, which is further evidnce for the non-linear property of tone production and perception. In sum, it seems fair to say that at least for some languages, that the same F_0 interval in different pitch ranges may not be articulatorily and perceptually equivalent.

In order to minimize the positive skew, the F_0 contour, if available, is normalized according to Zhu's (1999: 47) Logarithmic Z-score (henceforth LZ-score). Furthermore, tone levels will also be based on LZ-score, rather than on the raw F_0 value. This is mainly because we need to factor out inter-speaker variations and to draw out the invariance of a linguistic signal across a variety of contexts, e.g. citation and sandhi. Using the raw F_0 value alone is insufficient to meet these requirements. LZ-score transformations "[e]xpress an F_0 value as a multiple of a measure of dispersion from a mean value, all of which are in logarithmic terms" and is computed with the following formula.

$$z'_i = \frac{y_i - m_y}{s_y} = \frac{\log_{10} x_i - \frac{1}{n}\sum_{i=1}^{n} \log_{10} x_i}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n} (\log_{10} x_i - \frac{1}{n}\log_{10} x_i)^2}} \tag{29}$$

where $y_i = \log_{10} x_i$, x_i is an observed F_0 value, m_y and s_y are respectively the arithmetic mean and Standard Deviation of y_i. See Zhu (1999) for more discussion on how and why LZ-score achieves a better normalization result among various normalization approaches.

Finally, the above discussion also sheds light on the issue as to why the maximal contrasting tone levels are not more than five. This typologically robust generalization may be attributable to the (possible) logarithmic nature of the production and perception of tone. We have learned that pitch intervals tend to be larger in higher F_0 range. So it may well be the case that adding the sixth tone level in an inventory would lead to a pitch interval that is *too large to be implemented* from an articulation point of view. Let us consider Copala Trique (Figure 1-3) for example. The pitch range of the highest tone is around 30 Hz. If a sixth level were added to this inventory, its range would be at least 40 Hz. Given that the tone space in this language is approximately 75 Hz (Min: 85 Hz ~ Max: 160 Hz), a sixth level tone might enlarge the overall pitch range to at most 115 Hz. So we may entertain the possibility that the expanded space for the sixth level tone would contradict minimization of articulatory effort to a substantial extent, hence is cross-linguistically avoided. This speculation is of course subject to justification, but is nevertheless quite reasonable, as far as I can tell.

This completes our discussion on the logarithmic nature of tone production and perception. In the following section, I turn to the second little-discussed issue in the study of tone: contrasting pitch shapes.

1.3.2 Varieties of Contour Tones: Contrasting Pitch Shapes

> "I personally believe it is a serious reductionist mistake to try to force all the world's contours tones into a single analytical mould – James Matisoff (p. ix, Preface of *BLS 18: Special Session on the Typology of Tone Languages*)."

What Matisoff meant by "a single analytical mould" is the either/or opposition between the "atomic" and "molecular" approaches to a contour tone system in most, if not all, analyses of tone languages. The goal of this subsection is to show that in addition to the alleged "serious reductionist mistake," former treatments of contour shapes deserve a similar criticism: tonal contours are described in terms of the interpolation between beginning and end points which is too reductionist. This subsection deals with this problematic phenomenon for the standard model. That is, one of the oft-overlooked aspects of contour tone is that falling and rising tones do not always fall into a single category, i.e. a straight interpolation between two distinct tones, or pitch targets. Let us call them "*continuous* fall/rise" (which was termed "straight" fall/rise in Zhu (1999)). When seeing a transcription such as HL or LM, it is fair to say that we generally assume that phonetic realization of HL or LM is continuous fall/rise.

(30) The 'ideal' contour tones: Continuous fall/rise

Two exemplar continuous falls and rises are presented in Figure 1-8. For each example, waveform, pitch track (in Hz) and intensity contour (in dB) are presented. F_0 tracings and intensity contours of the tokens were made using Praat (Boersma and Weenink 2007). The diagrams in the left-hand column illustrate a continuous falling tone of the Hausa word *shâ* 'drinking' (data from the UCLA phonetics lab archive, http://archive.phonetics.ucla.edu/). In the right-hand column are the relevant phonetic data for the rising tone of the Lhasa Tibetan word *sǎ* 'land' (data from Zhou (1983). See also §5.4 for more details). As we can see, these two examples are alike in that the pitch contours are basically a straight interpolation between the starting point and endpoint of F_0.

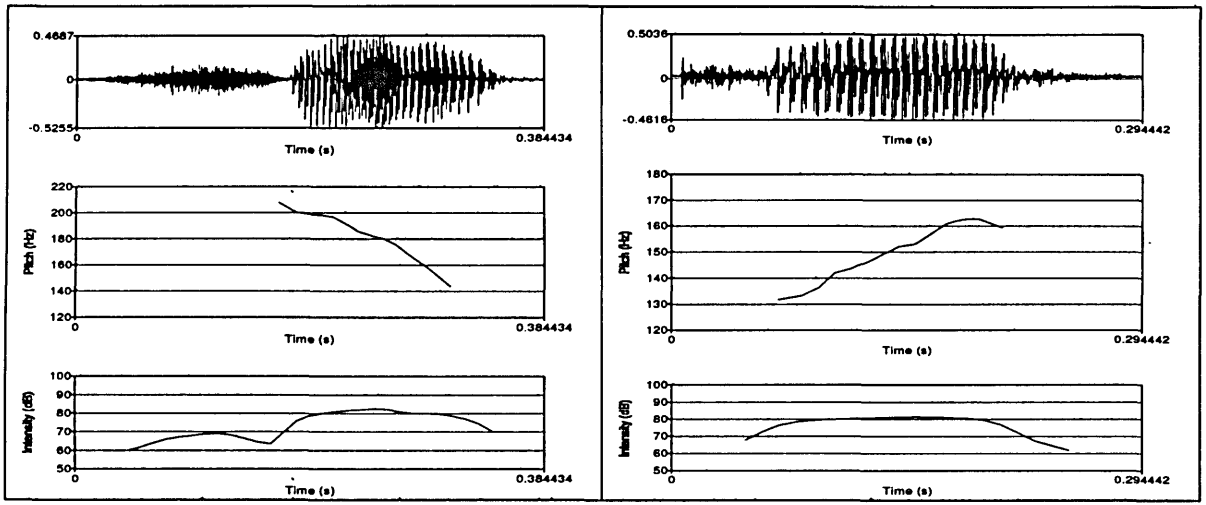


Figure 1-8 Exemplar F_0 contours (in Hz) and intensity (in dB) of continuous fall in Hausa (left) and continuous rise in Lhasa Tibetan (right)

Cross-linguistically speaking, continuous fall/rise is not the only attested shape of the falling and rising tones. It has been noted in Zhu (1999) and Yip (2001) that contour tones are not necessarily uniform in their shapes. First, it is attested that a plateau may appear in the initial portion of a contour tone. This type of contour tones was termed "delayed fall/rise" in Zhu (1999). Two exemplar contours of delayed fall/rise are in turn illustrated below. The delayed fall is from the Bangkok Thai word k^h*âː* 'I' (data from Esling 1994), and the delayed rise is from the Shanghai Chinese word *tie* (data from Ping (2004)).

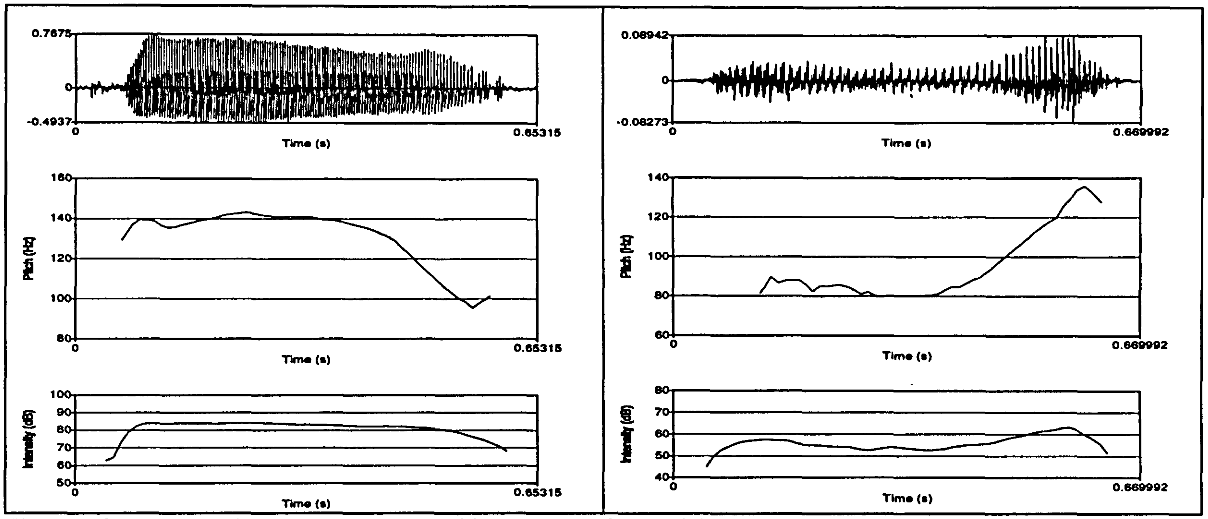


Figure 1-9 Exemplar F_0 contours (in Hz) and intensity (in dB) of delayed fall in Bangkok Thai (left) and delayed rise in Shanghai Chinese (right)

It can be seen in Figure 1-10 that an F_0 plateau may also occur in the final portion of a contour tone, which will termed "early fall/rise" in this work. The word *tuʌʔ* 'poison' illustrates an early fall and the word *huʌ* fire' shows an early rise in Pingyao Chinese (data from Qiao and Chen's (2004) *The Phonetic Database of Pingyao Chinese*).

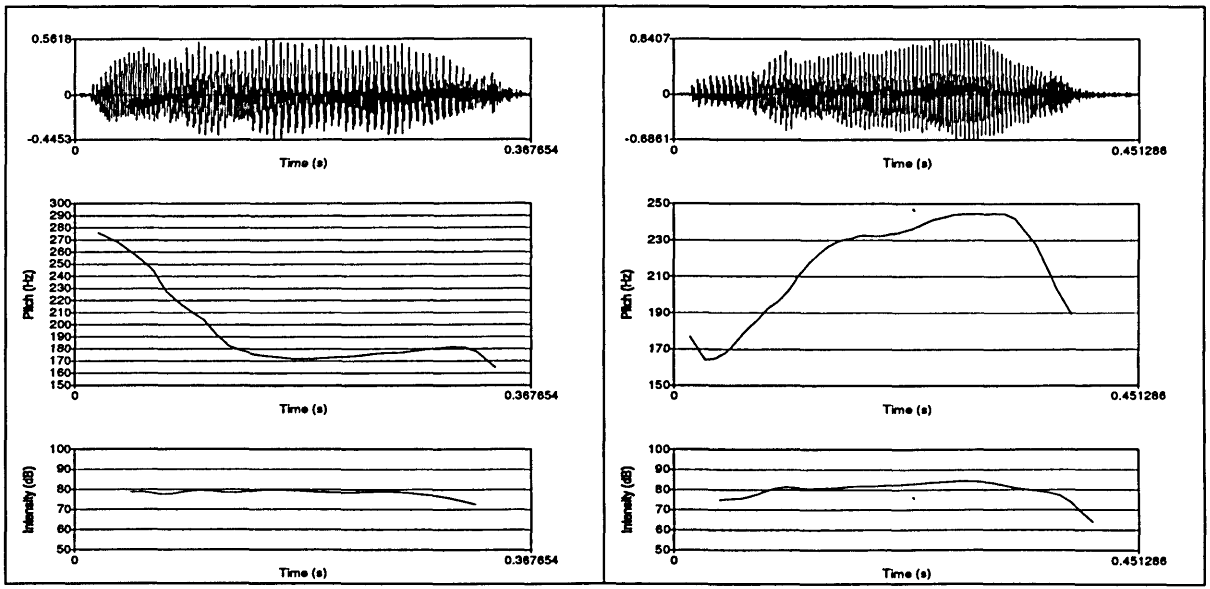


Figure 1-10 Early fall/rise in Pingyao Chinese

From the above illustration, it is obvious that contour tones do differ in their pitch shapes. An issue that arises here is how different pitch shapes should be analyzed. Given a binary approach, it is not clear how a continuous fall and an early fall can both be adequately represented with HL, or equivalent tonal features. Following SPE, pitch shapes might be regulated by some "low-level" phonetic rules. Based on this, one might contend that pitch shapes are never contrastive in a tone inventory. To my knowledge, this claim holds for most tonal systems. However, in Suzhou Chinese, a variety of Wu Chinese, continuous fall and early fall are both attested in the citation long tone inventory (i.e. tones on sonorant-final syllables). Observe now the F_0 contours of the citation long tones. The phonetic data are replotted from Sun's (2001) acoustic study on one male speaker. 9 tokens (3 monosyllabic words with 3 repetitions) were measured for each tone, using KAY CSL 4300B (see Sun (2001: 40) for the word list).

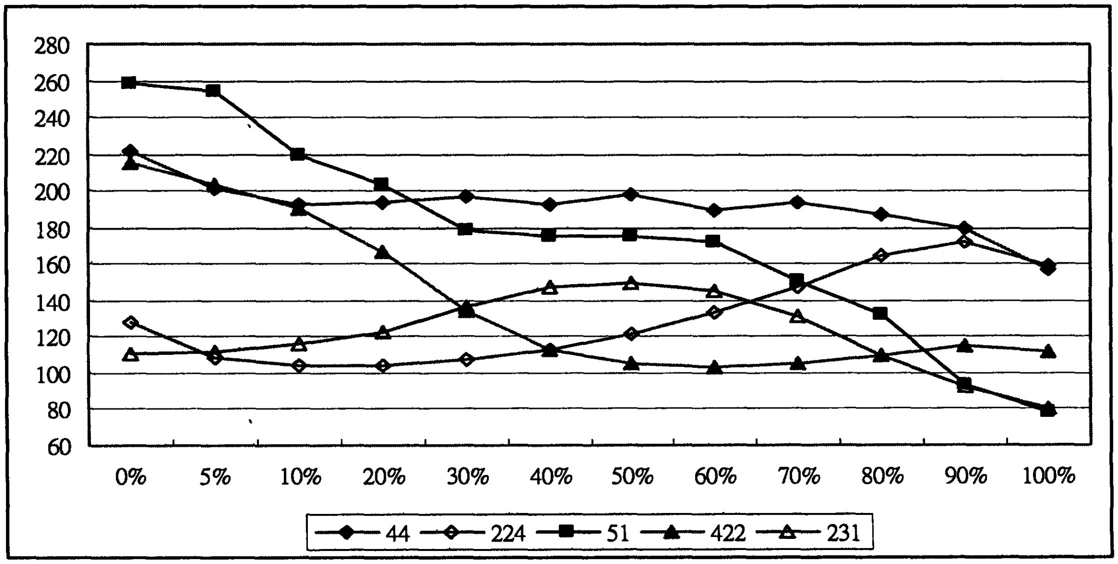


Figure 1-11 Normalized F_0 (in Hz) for non-checked citation tones in Suzhou Chinese

Suzhou Chinese has five non-checked tones. As a typical Wu Chinese language, tones are divided into the H-register (represented with the filled shape) and the L-register (represented with the hollow shape). It has been instrumentally confirmed that H-register tones only co-occur with modal phonation and L-register with breathy voice (or murmur) in Suzhou Chinese (Iwata et al. 1991). So the initial dip of the convex tone 231 can be analyzed as the effect of breathy voice.[17] In addition, absolute duration of each non-checked tone is presented below.

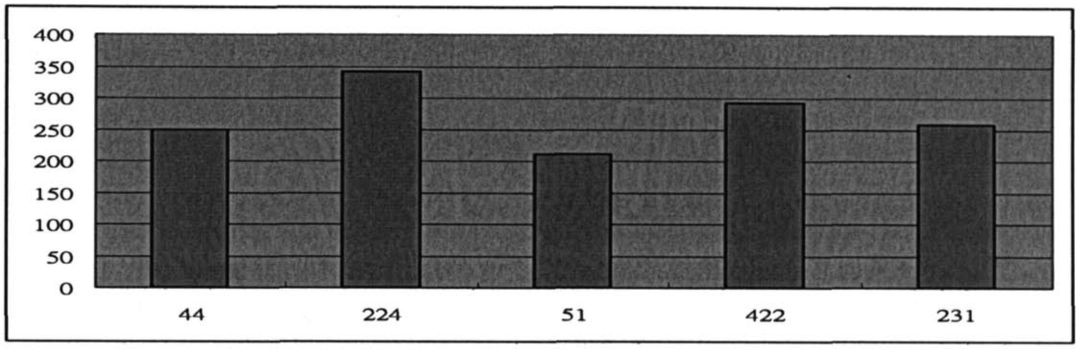


Figure 1-12 Duration of rime (in ms) for non-checked citation tones in Suzhou Chinese

Of present interest is the contrast between tones 51 and 422. An important difference between 51 and 422 lies in duration. As we can see in Figure 1-12, the continuous fall 51 (211 ms) is significantly shorter than the early fall 422 (292 ms) by 81 ms. For convenience, the pitch contours of the two tones in question are separately presented in Figure 1-13.

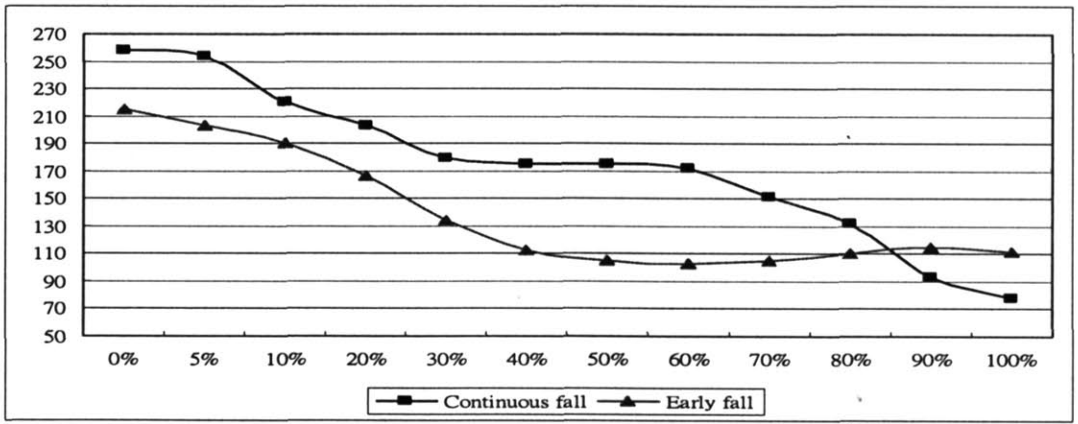


Figure 1-13 Normalized F_0 (in Hz) for continuous fall (51) and Early fall (422) in Suzhou Chinese

---

[17] See §2.3 and chapter 4 for more related discussion on the tone-consonant interaction in Wu Chinese. Recall that L-register tones do not always co-occur with murmur in Lishui Chinese, a Southern Wu Chinese language (Figure 1-1).

The final part of tone 422's contour is very flat. The pitch fluctuation is within a range of 10 Hz. We can see from Figure 1-11 that 10 Hz cannot be regarded as a phonologically contrastive pitch interval. If it were, there would be 18 tone levels in this language (i.e. Max: 260 Hz and Min: 80 Hz), contrary to fact. In other acoustic studies of Suzhou Chinese tonal system,[18] Liao (1994) transcribed tone 422 as 522, based on mean F_0 values from four speakers (2 males and 3 females), while in Lau's (2002) multi-speaker experiment (5 females and an unmentioned number of males), 422 was transcribed as 522, and he called it a "falling-level" tone.[19] Taken together, the results from these studies suggest that 422 is an early falling tone, that is, a falling tone with a plateau towards the end of the pitch contour, although the actual tone values are slightly different in each study, which is attributable to individual researcher's interpretation and analysis.

Returning to the representational issue, we can conclude that a binary approach is unable to handle tonal contrasts of this sort. Under the above-mentioned widely accepted assumption, contour tones are comprised of either two level tones or are specified with equivalent laryngeal features (modulo the convex and concave tones). Since the starting and ending pitches of the two tones are basically within the same ranges, in terms of Yip (1980, 1989) or Bao (1990, 1999), both early fall and continuous fall should be expected to have the same feature specification: [+U, HL], which cannot be the case. Thus, it should be evident that contrasting pitch shapes pose a non-trivial problem for the standard model.

In a gradient approach, on the other hand, it is not necessary to propose that a non-continuous rise/fall (i.e. delayed and early) has three pitch targets, e.g. 113 or 551. The

---

[18] In impressionistic transcriptions of Suzhou Chinese tonal system, tone 422 has been transcribed as 412, 513, or 523, probably due to dialect variation. See Shen (1995) for an overview.

[19] Lau (2002) also notes that the contrast between the continuous and early falls is being lost among younger generation speakers.

reason is the following. According to my acoustic study, based on Qiao and Chen's (2004) recordings of one male speaker, Pingyao Chinese has the following non-checked tone inventory: two early rises 244 and 144, and a concave tone 523, as the F_0 tracings below illustrate:

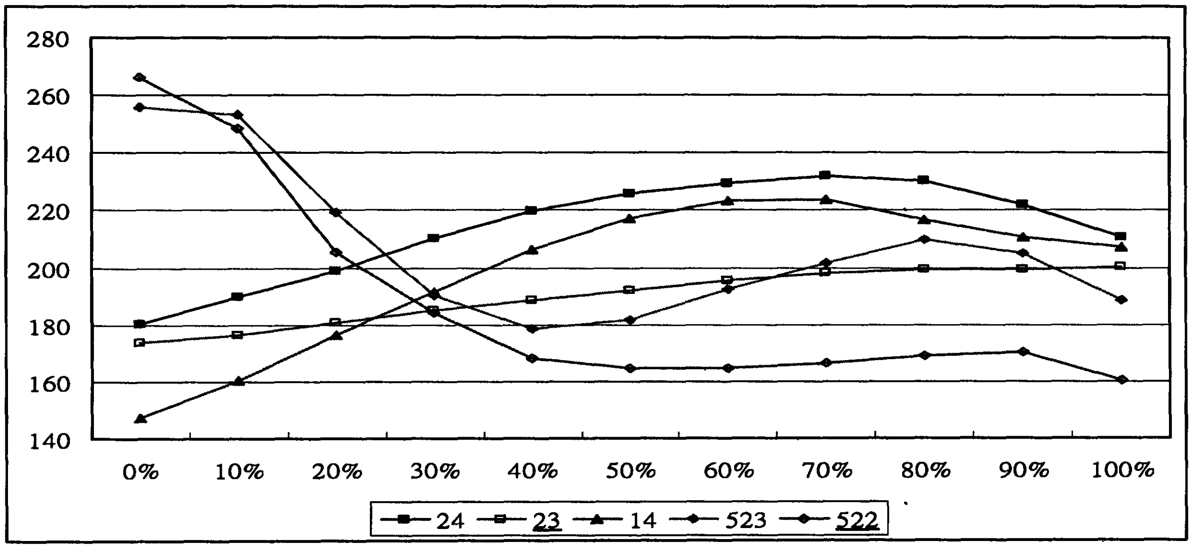


Figure 1-14 Normalized F_0 (in Hz) for citation tones in Pingyao Chinese

The non-checked tones can be all represented with three pitch targets: 244 (=24 in Figure 1-14), 144 (=14 in Figure 1-14), and 523. But this treatment will yield undesired results, e.g. the prediction is that there should be a tone inventory that only contains convex or concave tones.

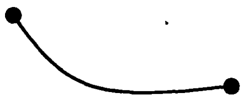
I suggest that a better way to avoid complications of this sort is to include *slope* in tonal specification. Let us consider the early fall (442) and continuous fall (51) in Suzhou Chinese (Figure 1-13). In (31), gross F_0 differences between the starting and ending points are provided, together with approximate rime duration. Recall from (18) in §1.2.4 that slope is defined as the following ratio: F_0 difference (the altitude change) over duration (the horizontal difference). In order to produce an early fall in 422, slope value must not be less than 0.6,[20] i.e. 100 Hz / 150 ms. If 422's slope value were specified as 0.3 (=100 Hz / 300 ms), then its contour shape would

---

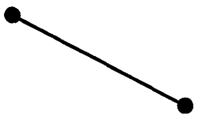
[20] Note that the slope value for a falling contour is negative; so the positive slope values used here are the absolute values. See §5.2.1 for more details.

be the same as 51's, i.e. a continuous fall, because there is no motivation for a sharp fall in the first half of the F_0 contour. By contrast, if 422's value is specified as 0.6, then the early rise would be induced to maintain the F_0 difference constant: 100 Hz. Otherwise, to fully realize the slope value 0.6 over duration of 300 Hz would lead to an undesired outcome: excessive pitch excursion: 180 Hz (=300 ms*0.6).

(31) Suzhou Chinese: 422 vs. 51



422:
Slope *specified as 0.6*
$\Delta F_0 \approx 100$ Hz
Duration ≈ 300 ms



51:
Slope specified as 0.75
$\Delta F_0 \approx 150$ Hz
Duration ≈ 200 ms

Above is a sample demonstration showing that it is beneficial to include slope in tonal specification of an early fall. Likewise, early rise and delayed rise/fall can be generated in a similar fashion.

To sum up, the main goal of this subsection is to show that pitch shapes are indeed phonologically contrastive. This specific property again cannot be easily handled in the standard model. I have suggested that a gradient approach with the specification of slope better accommodates the empirical facts.

1.3.3 Summary of this Section

In this section, I have touched on two little-discussed aspects of tone, namely the (possible) logarithmic nature of tone production and perception as well as contrasting pitch shapes. I hope that it is now clear that the standard model of tonal representation is too restricted to account for the attested tone inventories. According to the above discussion, I adopt a scalar representation of tone, i.e. multi-valued features on a five-point scale, (Chao 1930) throughout this work. Finally, under the assumption that the acoustic and auditory representations of tone are based on LZ-normalized F_0, rather than the raw F_0 value, contrasting tones are derived by a ranking of MINDIST and MAXIMIZE CONTRASTS constraints (Flemming 1995, 2002, 2006).

1.4 Organization of this Dissertation

I have observed above that relational correspondence has two constraint families, i.e. contour and slope correspondence. Chapter 2 introduces the formalization of the contour correspondence constraints, together with evidence that the unit correspondence and markedness constraints are unable to handle. The following two chapters apply contour correspondence to the analysis of tone sandhi phenomena in genetically diverse languages. In chapter 3, I propose that contour correspondence should be contextualized with respect to different proximities. Chapter 4 presents an analysis of the (non-)contouricity agreement phenomenon in Shaoxing Chinese. Slope correspondence is addressed in chapter 5, and conclusions are presented in chapter 6.

Chapter 2 Fundamentals of Contour Correspondence

2.1 Introduction

The previous chapter was an attempt to motivate relational correspondence. With regard to the auditory dimension F_0, I have proposed that relational correspondence has two distinct yet closely related constraint families, namely, contour correspondence and slope correspondence. In this chapter, I will formally define the contour correspondence constraints. The bulk of this chapter is then devoted to providing support for contour correspondence, in particular, preservation of the F_0 contour in the (syllable) *nucleus*.

The idea of contour preservation is far from new. It has antecedents in the OT literature. For example, Yip (2001) speculates in her discussion of Barasana that its tone-accent interaction makes "crucial reference to tonal sequences." Subsequently, Yip (2002: 197) proposes PRES-LH "to preserve the LH rise intact on a single syllable." A similar idea is expressed in Zhang's (2002b: 122) REALIZE-$\overset{\frown}{\text{HL}}$ "realize the $\overset{\frown}{\text{HL}}$ contour in some fashion." Finally, Zhang's (2006: 242) IDENT-TT(R) says that "the two adjacent rising contours in the input should also have two tone types in the output (no level tone)." While these constraints are formulated to cope with ostensibly different tone sandhi phenomena, the intuition behind these proposals should be obvious, i.e. preservation of the phonological height relation between successive tones. As we have discussed in chapter 1, however, contour preservation cannot be adequately expressed in the standard theory of faithfulness, i.e. unit correspondence. So it is fair to say that the above-mentioned proposals are non-standard in that the element-based infrastructure for unit correspondence is circumvented.

If we were to adhere to the orthodox faithfulness constraints, it would be conceivable that there might be two strategies to achieve the effect of contour preservation. First, we could resort to constraint conjunction (Smolensky 1993, et seq. and many others), e.g. MAX-(H)&MAX-(L) or IDENT-(L)&IDENT-(H), although we have seen in §1.2.2 that conjoining faithfulness constraints fails to yield the desired results in the English loanwords into Yoruba. This approach was implicitly adopted in Yip's (2002) PRES-LH and Zhang's (2002b) REALIZE-$\widehat{\text{HL}}$ constraints. Second, employing more sophisticated tonal representation would work. In particular, a H-register rising tone may be realized as a L-register rising tone by changing the underlying register feature [+U(pper)] to [-U(pper)] on the surface, while tonal (contour) specifications are faithfully realized: [+U, LH] → [-U, LH].[1] In summary, we might be led to the conclusion that contour preservation can be handled equally well given unit correspondence. To see how contour correspondence is distinguished from these alternatives, as the first step, I will discuss the essential assumptions of contour correspondence in the auditory dimension F_0; then I will develop a formalization of the contour correspondence constraints in terms of Optimality Theory (Prince and Smolensky 1993, 2004).

2.2 Defining (Nucleus-internal) Contour correspondence

The general schema of contour correspondence is essentially couched in Steriade's (2006) "contour preservation" constraints:

---

[1] In addition to this, it is possible to "reintroduce" the atomic contour features such as [±FALLING] or [±RISING] proposed in Wang (1967). Then, we may posit IDENT-[+RISING] or MAX-[+FALLING] to enforce contour preservation, if these constraints are active in deciding the winner. However, feature systems of this sort usually, if not always, predict unattested phenomena. For example, spreading [-RISING] to the following syllable will result in a non-rising tone, i.e. either level or falling tone, which does not seem to be attested. Furthermore, if we postulate that atomic contour features are privative, this formulation fails to capture some of the most well-attested tone spreading rules, e.g. L.H → L.R, in which [RISING] must be arbitrarily inserted.

For proposals along this line, see Clark's (1978) "downward arrow" that marks the pitch drop as a unit and Xu and Wang's (2001) linear pitch target.

(1) Contour preservation constraints (Steriade 2006)[2]

D, an auditory dimension
xy, a sequence of elements in S_1, where x precedes y
ab, a sequence of elements in S_2, where a precedes b
x is the S_1 correspondent of a
D(x) = the D value of x; D(a) = the D value of a; etc.

a. NO REVERSAL

If $(D(x) < D(y))$ then $\neg(D(a) > D(b))$ and
if $(D(x) > D(y))$ then $\neg(D(a) < D(b))$

b. MATCH

If $(D(x) < D(y))$ then $(D(a) < D(b))$ and
if $(D(x) > D(y))$ then $(D(a) > D(b))$ and
if $(D(x) = D(y))$ then $(D(a) = D(b))$

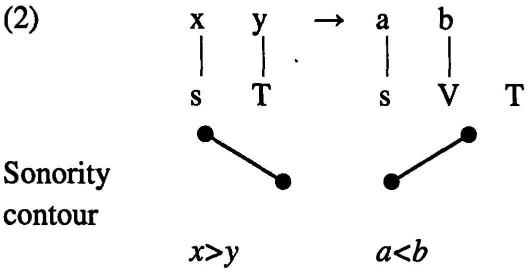
In essence, contour preservation is *syntagmatic* contrast preservation, i.e. contrast maintenance between two successive elements, as opposed to paradigmatic contrast according to which the well-formedness of an entity is evaluated with reference to the set of entities it contrasts with (Flemming 1995, 2002, 2006). Syntagmatic contrast is manifested in various auditory dimensions. Steriade (2006) suggests that auditory domain *D* includes i) intensity/loudness (sonority), ii) metrical prominence, and iii) F_0 (pitch). To give an example, we have discussed in §1.1 that contour preservation provides a more straightforward account for the different separability of obstruent-sonorant (TR) vs. fricative-obstruent (sT) clusters. This asymmetry lies in the distinct sonority contours of TR vs. sT: TR exhibits an underlying rising contour and sT an underlying falling contour. Assuming that "element" here refers to the degree of the loudness a segment bears, Steriade postulates that the rising/falling sonority contour across two segments *xy* is represented with "$x<y$"/"$x>y$" because *x* is perceptually less loud/louder than *y*. In other words,

---

[2] Steriade's (2006) SLOPE constraint will be discussed separately in chapter 5.

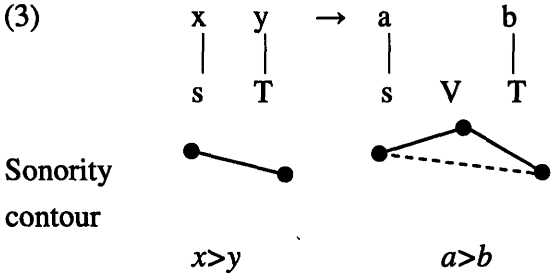
the contour is interpreted as the loudness relation between two successive elements, e.g. *x>y* or *x<y*. A particular relation should be maintained on the surface if contour preservation constraints are active: *x>y* maps to *a>b*, but not *a<b*, given that *a* is the correspondent of *x*. Notice that *b* is *not* necessarily the correspondent of *y*. This is because when a consonant cluster is broken by vowel epenthesis, for example, sT → sV(T), No Reversal in (1)a is violated in the mapping below.

(2)



Let us now consider what happens if *b* is the correspondent of *y*. Given that *xy* = sT and *ab* = sT, sT → s(V)T will not incur a violation of either No Reversal or Match since the relation between sT remains the same as long as *s* precedes *T*: the solid line between *xy* and the dashed line between *ab* exhibit the same relation: "greater than."

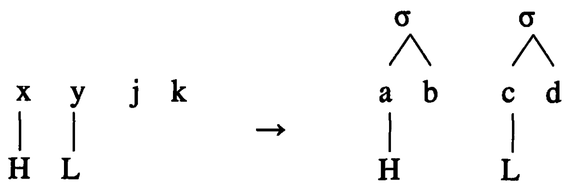
(3)



Under the current formulation, precedence is not defined as *immediate* precedence (cf. De Lacy 2007); consequently the intervening epenthetic vowel does not play a role in this regard. This problem, as Steriade suggests, can be fixed by leaving the following corresponding elements undefined (i.e. *b* in (1); see also (2)).

Turning back to tone, it is well-known that the characteristics of tone are substantially different from other auditory dimensions such as loudness (sonority). For a better understanding, consider the following schematic example. Suppose that *xyjk* in the input and *abcd* in the output are corresponding tone-bearing units and L de-links from *y*, re-associating to *c*, as in (4) below. As we can see, this example illustrates the following specific properties of tone. *Tone stability* and *tone mobility* discussed in §1.2.1 and §1.2.4 are rarely observed in segmental features (and their auditory correlates). More importantly, there does not appear to be a comparable instantiation of *tonelessness*[3] in consonantal and vocalic features. If the relation between successive tones is defined within the (syllable) nucleus, the mapping below is not faithful because the surface HL contour does not remain intact on a syllable. On the other hand, the HL contour is nonetheless preserved across syllables in the output representation. It should be obvious that sonority does not work in a similar fashion. For example, sonority is never deleted from a segment and re-linked to an adjacent host. Moreover, every segment must bear some degree of loudness or sonority, as opposed to tonelessness. Therefore, it appears that the mapping in (4), or comparable cases, is attested only in tone.

(4) Why tone is different: Tone mobility, tone preservation, and tonelessness



---

[3] Notice that tonelessness here should be understood as a tone whose surface pitch is acquired by phonetic interpolation, not a phonologically inert tone.

For this and other reasons (to which I will return in chapter 5 where slope correspondence is introduced), we need to postulate a formulation of contour correspondence that is specific to the F_0 contour. As the starting point, I adopt Steriade's (2006) schema according to which there are two essential components in the formulation of contour correspondence, as stated below.

(5) a. What is a possible relation of the F_0 contour?
    b. What "elements" stand in contour correspondence in tone?

We will consider these components of contour correspondence in turn in the section that follows. Also, from now on, I will refer to "contour correspondence in tone and tone sandhi" as Contour correspondence throughout the discussion.

2.2.1 Four Relations between Successive Tones within the Nucleus

In this section, I address the issue as to what a possible relation in the F_0 contour is. In the auditory dimension F_0, "relation" refers to the phonological height or scaling relation between successive tones. Supposing that tone T_1 immediately precedes tone T_2 in a (syllable) nucleus, I propose that there are four distinct height relations.[4]

---

[4] The present proposal resembles "relational features" in Dilley's (2004) tone interval theory in many respects. Relational features specify the abstract spatial configuration between a *referent* tone and a *referring* tone in a tone interval construct (roughly speaking, a minimal bitonal unit). She proposes the features [±same] and [±higher] to derive three relative height relations between two tones: *higher, lower* and *same.* For present purposes, it is important to note that a substantial distinction between the two approaches lies in the fatc that the four relations listed in (6) are *not* specifications *per se.* In contour correspondence, maintaining the relation (or the contour) between two tones in the input and the output is regulated via correspondence. That is, the relation in the input and in the output should be as similar as possible. More discussion of different predictions of the two approaches is postponed to §2.4 and §3.4.

(6) Four relations between two successive tones *within the nucleus* (first approximation)

a. The "greater than" relation = Falling tone
b. The "less than" relation = Rising tone
c. The "non-equal to" relation = Falling or Rising tones
d. The "equal to" relation = Level tone

The first two relations should be straightforward. Falling tone constitutes the "greater than" relation because H's tone level is higher than L's. Likewise, LH, or a rising tone, form the "less than" relation, because L's tone level is lower than H's. As for the "non-equal to" relation, this can be instantiated by any two unlike tones. For example, either HL or LH constitutes the "non-equal to" relation since the preceding tone is not equivalent to the following tone. In sum, the three relations we discussed so far all involve the relation between two unlike tones.

The final relation needs more discussion. From the preceding discussion, it should be the case that the "equal to" relation within the nucleus refers to *two successive identical tones*, as in (7)b. As is well-known, (7)b is ruled out by the Twin Sister Convention (Goldsmith 1976, Clements and Keyser 1983, Odden 1986).

(7) How do we relate the "equal to" relation to a single tone?

a.

b.

The Twin Sister Convention is motivated by the following reasons. First, the two structures in (7) are auditorily indistinguishable. Furthermore, they also convey the same articulatory instruction. Autosegmental orthodoxy says that contour tones are comprised of at least two distinct tones, while it seems unnecessary for level tones to have two identical tones. Second, more importantly, (7)b is subject to the OCP, banning two adjacent identical elements on a melodic tier (McCarthy 1986). There is no denying that the OCP can be regarded as a violable constraint but one of the consequences is that there should exist contour tone-only languages if the OCP is top-ranked, ruling out HH, LL, etc, within a syllable (nucleus). To my knowledge, no tonal system consists of only contour tones. Notice that by "tonal system," I mean *both* the citation and the sandhi tone inventories.[5] Therefore, it is difficult to explain the rarity or non-existence of contour tone-only languages given the assumption according to which *all* tone (on a syllable) is comprised of a sequence of at least two tones. In sum, we are led to the conclusion that level tones should comprise only a single tone.

Since the general assumption is that level tones have only a single tone, as in (7)a, it then turns out that the preceding discussion calls for a reconsideration of what "elements" should stand in contour correspondence. More specifically, if level tones have only one single tone, how is the "equal to" relation established?

2.2.2 Constraints on (Nucleus-internal) Contour Preservation

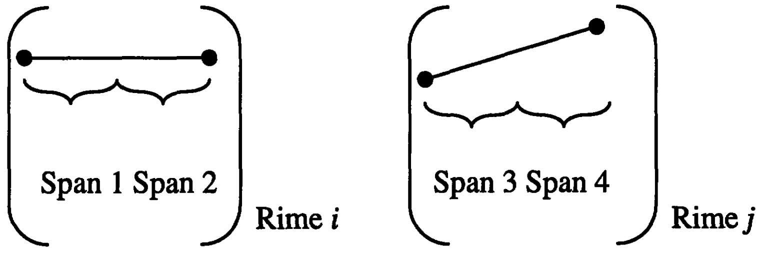
Since it seems problematic to define the "elements" standing in contour correspondence as "tones," I propose that these "elements" are "temporal spans," or "slices," of the F_0 contour. Before moving on to the definition of "slice," we need to clarify how the F_0 value is available in the underlying representation. Recall from §1.3 that I have presented empirical data in support of the claim that tones are better represented in the five-point scale, while pitch levels are expressed in (Logarithmic Z-score (LZ)-normalized) F_0 in Hz (cf. §1.3.1), instead of with the "standard" approach according to which the representation of tone is based on a restricted set of distinctive

---

[5] According to Zhang (1999), there is no level tone in the citation and the sandhi tone inventories in Pingyao Chinese. My unpublished acoustic study, based on Qian and Chen's (2004) recordings, shows that level tones are attested in sandhi context. This disparity might be due to dialect variation.

features. Given this assumption, the F_0 contour is specified in the input.[6] Returning to the definition of the temporal span, I hypothesize that the F_0 contour contained within a sonorous rime can be divided into at least two temporal spans. This stipulation can be motivated by the following tendency: the F_0 maxima and minima usually occur on the edges of a tonal contour.[7] So it is reasonable to posit the "two-span" construct. The idea is graphically illustrated below. I give two schematic examples, level M tone and rising tone L͡H. Temporal spans of falling tones are determined in the same way.

(8) Temporal spans



a. Level tone M

b. Rising tone L͡H

For clarity, two points must be addressed. First, it is important to note that temporal span is NOT a tone-bearing unit such as the mora, or more generally, a phonological entity. For example, level tones on a monomoraic syllable have two temporal spans, too. Recall that the central thesis of

---

[6] I assume that all phonetic details are specified in the input. This assumption is reminiscent of the hypothesis of *inferred input*, i.e. a hypothesized phonetic interpretation of the input (Steriade 1997, Jun 2002). Building on McCarthy's (1999) formalism of Sympathy Theory, Jun (2002: 13) proposes that "the inferred input be the most harmonic among candidates which obey ALL context-free IO faithfulness" (small capitals in original). See also the Fully Faithful Candidate (FFC) in McCarthy's (2003a) Comparative Markedness theory.

In Flemming's (2006) inventory-based model, the Realization component maps a string of segments from the inventory onto its phonetic realization. In particular, the assignment of the F_0 value takes place here, too. According to Flemming's formulation, markedness constraints apply in this component as well, indicating that the F_0 contour may be altered. In this work, I assume that the tone inventory is derived by a ranking of MINDIST and MAXIMIZE CONTRASTS constraints.

[7] We have discussed in §1.3.2. that slope should be specified in non-continuous contour tones (i.e. delayed and early rise/fall), rather than adding an extra tone, e.g. HHL. So in terms of temporal span, I assume that tones of this sort have two spans, too. As for convex and concave tones, three temporal spans may be needed.

this dissertation is that phonological computations must consider relations between successive elements. Temporal span is introduced as an analog of a listener's perceptual process when assessing a heard tonal contour to a phonological height, or scaling relation. Second, notice that the F_0 contour is contained *within a sonorous rime*. The reason is the following. It has been extensively discussed in the literature, especially in Gordon (2001) and Zhang (2002b), that sonorous rimes (including vowel and/or vowel-sonorant combinations) are better tone carriers than obstruent codas.[8] So it is reasonable to posit that only on the sonorous rime is the pitch contour linguistically relevant. Taken together, the preceding discussion makes it possible to formulate the constraint family of contour correspondence. In order to emphasize the central idea according to which it is the phonological height *relation* that is preserved, I will term this constraint family RELATIONAL CORRESPONDENCE (abbr. RELCORR).

(9) RELATIONAL CORRESPONDENCE (RELCORR)

Let t_1 be a tone value contained within Rime R. Let S_1 be a temporal span associated with t_1
Let t_2 be a tone value contained within Rime R. Let S_2 be a temporal span associated with t_2.
S_1 precedes S_2.

Let t_1' be the correspondent of t_1 in Rime R' and S_1', the temporal span associated with t_1' is the correspondent of S_1.
Let t_2' be the correspondent of t_2 in Rime R' and S_2', the temporal span associated with t_2' is the correspondent of S_2.
S_1' precedes S_2'.

Let $t_1 = x$, $t_2 = y$, $t_1' = a$, and $t_2' = b$.

i. The "greater than" relation: If $x > y$, then $a > b$. (abbr. RELCORR($x > y$))
ii. The "less than" relation: If $x < y$, then $a < b$. (abbr. RELCORR($x < y$))
iii. The "equal to" relation: If $x = y$, then $a = b$. (abbr. RELCORR($x = y$))
iv. The "non-equal to" relation: If $x \neq y$, then $a \neq b$. (abbr. RELCORR($x \neq y$))

---

[8] Let us ignore the tone-bearing ability of onset consonants for now, assuming that they are also worse tone bearer. I shall return to this issue in chapter 3.

There are several things to note about the definition of RELCORR. First, "tone value" refers to the tonal specification corresponding to a temporal span. It is conceivable that, for example, a 1 Hz difference is too small to be discernible so that the "greater than" relation cannot be accordingly established.[9] However, using the raw F_0 difference alone is difficult to manage, for instance, inter-speaker variations. More importantly, the degree of the raw F_0 difference can only be perceived as contrastive by phonological computations. So I use the linguistic term tone (value), rather than the physical term "F_0", or the perceptual term "pitch" in (9). Then, assuming that the tone inventory is derived by a ranking of the MINDIST and MAXIMIZE CONTRASTS constraints (Flemming 1995, 2002, 2006), the proposed four relations are established by reference to contrasting tones. More precisely, the tone value associated with temporal span *x* is, for example, greater than that associated with temporal span *y*, if and only if *x*'s tone value contrasts with *y*'s and *x*'s tone value is higher than *y*'s in terms of F_0. The other relations are derived in the same fashion.

Second, note also that segmental affiliation is *not* included in the formulation of RELCORR. Rime R' is not necessarily the correspondent of Rime R in (9). This is due to the characteristics of tone, e.g. tone mobility, tone preservation, etc: it is well-known that a tone may be realized on a syllable that is not its lexical source. I propose that the IDENT-(Tone) constraint is responsible for maintaining the exact identity of the tonal specification of a segment, whose tone-bearing ability is determined on a language-specific basis (see Gordon 2001, Zhang 2002b for different proposals).

(10) IDENT-(Tone) (abbr. IDENT-(T))
"Correspondent segments have identical values for the tonal specification."

---

[9] The minimal detectable difference for an F_0 ramp is 2 Hz (Klatt 1973).

One of the consequences of a lowly ranked, inactive IDENT-T is that tone is free to migrate from its underlying host to another host given that vowel deletion, glide formation, or tone attraction takes place.

Third, it is also important to note that RELCORR is indexed to the four relations and the indexed RELCORR constraints are freely rankable with one another. This approach is essentially different from NO REVERSAL and MATCH in (1) (repeated in (11) below), in that Steriade's (2006) contour preservations constraints are "monolithic." That is, contour preservation requires that *all* types of contour (rise/fall/level) should be faithfully realized on the surface. In particular, MATCH is violated, for example, by a rising-contour-to-level-contour mapping, even though the other types, i.e. falling and level contours, remain intact in the output.

(11) "Monolithic" contour preservation constraints (Steriade 2006)

D, an auditory dimension
xy, a sequence of elements in S_1, where x precedes y
ab, a sequence of elements in S_2, where a precedes b
x is the S_1 correspondent of a
$D(x)$ = the D value of x; $D(a)$ = the D value of a; etc.

a. NO REVERSAL

If $(D(x) < D(y))$ then $\neg(D(a) > D(b))$ and
if $(D(x) > D(y))$ then $\neg(D(a) < D(b))$

b. MATCH

If $(D(x) < D(y))$ then $(D(a) < D(b))$ and
If $(D(x) > D(y))$ then $(D(a) > D(b))$ and
If $(D(x) = D(y))$ then $(D(a) = D(b))$

While NO REVERSAL or MATCH may be sufficient for sonority contour, such monolithic contour preservation constraints are not suitable for the analysis of tonal contours. To give an example, faithful realization of rising tone does not necessarily entail that level tone must remain level on the surface. This asymmetry is documented in Comaltepec Chinantec (Silverman 1997 and references cited therein). In brief, this language has the following phonotactic constraint: *L͡H.L, where syllable boundary is marked with a dot. Consider now the following pair.

(12) Comaltepec Chinantec: Contour preservation is not always enforced across the board

a. L͡H.L → L͡H.**H͡L**
b. L͡H.L͡M → L͡H.L͡M

If a low level tone L is preceded by a high rising tone L͡H, L becomes H͡L in order to satisfy the markedness constraint *L͡H.L, indicating that the underlying contour on the following syllable in (12)a is altered: from the input level contour (x=y) to the output falling contour (a>b). In (12)b, by contrast, L͡M remains L͡M in the same environment. For the present purpose, we can say that if the rising contour of L͡M were protected by the ranking MATCH » *L͡H.L, then there would not appear to be good reason as to why the low level tone should change to a falling tone to avoid violation of *L͡H.L. Given the monolithic view of contour preservation in (11), *all* types of contours are supposed to be equally shielded by the active MATCH constraint. However, as we have seen, this is not the case, as far as the tonal contour is concerned. Consequently, I take the stance that each of the four relations is indexed to RELCORR and these indexed RELCORR constraints can be freely ranked with one another. For the case at hand, the following ranking yields the desired results: RELCORR (x<y) » *L͡H.L » RELCORR (x=y). Informally, preservation of the rising contour is "more important" than that of the level contour, under pressure from the same markedness constraint. More discussion of an analysis along this line will be provided in many examples in subsequent chapters.

This completes our introduction to the contour preservation constraint family, RELCORR. To recapitulate, the essential assumptions made in this model are enumerated as follows.

i. Tones are represented in terms of the five-point scale, while pitch levels are expressed in (LZ-normalized) F_0 in Hz. The tone inventory is derived by a ranking of the MINDIST and MAXIMIZE CONTRASTS constraints (Flemming 1995, 2002, 2006).

ii. The hypothetical "temporal spans" of a tonal contour (contained within a rime) stand in contour correspondence, not tone *per se*. In addition, temporal span is *not* a tone-bearing unit (e.g. the mora).

iii. There are four phonological height or scaling relations between tone values associated with two successive temporal spans: i) greater than, ii) less than, iii) equal to, and iv) non-equal to.

iv. The contour correspondence constraint family (RELCORR) assesses relational similarity of tone values associated with two successive temporal spans.

v. At least for tone, I posit that the four relations in (iii) are indexed to RELCORR, so that the indexed RELCORR constraints are freely rankable with one another, e.g. RELCORR ($x<y$) » RELCORR ($x>y$), etc.

2.2.3 Plot of the Analysis

So far I have laid out the general schema of contour correspondence and formally defined the contour correspondence constraint family in tone sandhi: RELCORR. The most straightforward evidence for contour correspondence comes from the phenomenon in which the contour is preserved, regardless of the faithfulness violations of the individual input tones. In this chapter, we confine our attention to such phenomena in the (syllable) nucleus. The contour correspondence proposals are developed in one case study. This case study concerns the phenomenon I termed "Invariance of Variation" in §1.2.3, i.e. the overall F_0 value of Tone T is changed but the rising/falling/level profile is not. For example, the checked rising tone in Shanghai Chinese features the following mapping: "L-register rising tone → H-register rising tone" in disyllabic tone sandhi. I will argue in §2.3 that the Shanghai Chinese data cannot be adequately handled with the known analytical tools, in particular, unit correspondence constraints and with the existing representational devices, i.e. the multi-planar tonal nodes (i.e. the register and contour tone nodes). In addition to this, recall that we have postulated four scaling relations in contour correspondence. We have also briefly discussed that RELCORR should be indexed with the four relations and the indexed RELCORR constraints are freely ranked with one another. In §2.4, I preview the case studies for the proposed four scaling relations in subsequent chapters, in particular, the "greater than," "equal to," and "non-equal to" relations. Finally, this chapter is closed by a discussion of comparisons with alternative approaches to contour preservation.

2.3 Shanghai Chinese: Invariance of Variation in Register-raising

The arguments for contour correspondence constraints hinge on the existence of the phenomena in which the desired result cannot be obtained by unit correspondence, e.g. MAX, IDENT, or constraints that preserve or approximately match the F_0 value of individual input tones. We have briefly sketched the point in §1.2.3, i.e. the "Invariance of Variation" phenomenon. As a reminder, consider first the following mapping. Tone notation is based on the five-point scale, according to which 5 is the highest and 1 the lowest.

(13) 23 [-U(pper), LH] → 34 [+U(pper), LH]

In this mapping, the L-register rising tone 23 raises to the H-register rising tone 34. It appears that tonal identity is not exact with respect to each pair of the input and output tones. More precisely, on the surface, 2 ([-U, L]) becomes 3 ([-U, H] or [+U, L]), while 3 ([-U, H] or [+U, L]) changes to 4 ([+U, H]). It appears that IDENT-(Tone) is not satisfied because tonal specifications are not identical between the input and the output. MAX-(Tone) is violated, too, since 2 and 3 are not faithfully realized in the output. Finally, the surface tones 3 and 4 do not have their input correspondents, incurring violations of DEP-(Tone). At first sight, it seems that *register-raising* (i.e. a mapping from an input L-register tone to an output H-register tone) cannot be adequately expressed by unit correspondence constraints. Element-based correspondence is, however, not entirely hopeless in the case at hand. It is conceivable that register-raising could be generated by change of register feature specification on the register node (e.g. Yip 1989, Bao 1990, 1999). For the present purpose, let us simply stipulate that tone level 3 can be optionally represented with either [+U, L] or [-U, H]. This treatment was termed the "dual structure of the mid tone" in Bao (1990, 1999). See also Yip (2001) for a dispersion-theoretic approach.

(14) 23 → 34

[-U, LH] [+U, LH]

Employing the multi-planar tonal node seems to capture the phenomenon in question in a straightforward fashion. However, it is important to note that this analysis would work if and only if insertion of the register feature is not penalized by the active anti-insertion constraint DEP. This point is illustrated in the following tableau.

(15) Register-raising is banned if insertion of [+U] is disallowed

[-U, LH]	*[-U, LH]	Dep-(+U)	Max-(-U)	Max-(H)	Max-(L)
a. [-U, LH]	*!				
b. [+U, LH]		*!			
☞ c. [-U, L]				*	
☞ d. [-U, H]					*

Suppose that Dep-(+U(pper)) is top-ranked, indicating that insertion of [+U] is not an option to avoid violation of the undominated markedness constraint *[-U, LH]. It can be seen in (15) that under the pressure from [+U, LH], Dep-(+U) eliminates candidate (b), the output form featuring register-raising, while contour flattening is the optimal strategy to satisfy *[-U, LH], as shown in candidates (c) and (d). So it would be surprising that the rising contour is nevertheless retained under the current ranking of constraints. As we will see in the following sections, this possibility is indeed attested in Shanghai Chinese. Our analysis can be sketched as follows. Provided that the contour correspondence constraint RelCorr ($x<y$) is top-ranked, it requires that the rising contour should be preserved on the surface, even at the expense of violation of Dep-(+U).

(16) Register-raising is protected by RelCorr

[-U, LH]	*[-U, LH]	RelCorr ($x<y$)	Dep-(+U)
a. [-U, LH]	*!		
☞ b. [+U, LH]			*
c. [-U, L]		*!	
d. [-U, H]		*!	

I will show in the following section that register-raising of the Shanghai Chinese checked rising tone in sandhi approximates this specific phenomenon, hence as crucial evidence for contour correspondence.

2.3.1 The Citation Tone Inventory in Shanghai Chinese

Shanghai Chinese is one of the representative dialects of Wu Chinese. The variety of Shanghai Chinese under discussion is New Shanghai Chinese (henceforth SH). There is a substantial body of research on tone and tone sandhi in this language (see Chen 2000 for references cited therein). This language has five tones in the citation tone inventory: three (non-checked) long tones and two checked tones. Unless otherwise noted, data are drawn from Ping's (2001b) acoustic study.[10]

(17) The citation tone inventory of Shanghai Chinese

a. Long tones (i.e. tones on sonorant-final syllables)

H-register	Rime	SD	N	L-register	Rime	SD	N
Tone 1: 551	185.9	27.2	9	Tone 3: 113	242.0	24.3	9
Tone 2: 445	243.5	21.1	9				

b. Checked tones (i.e. tones on glottal stop-terminating syllables)

H-register	Rime	SD	N	L-register	Rime	SD	N
Tone 4: 44	92.4	13.0	9	Tone 5: 13	134.0	12.9	15

(Pitch values are transcribed in the five-point scale (i.e. Chao's letters), where 5 is the highest and 1 the lowest; Rime duration is given in ms; SD=Standard Deviation; N=number of tokens; checked tones are underlined throughout.)

One of the most distinctive traits in Wu Chinese is the "register" contrasts in the citation tone inventory. It has been instrumentally confirmed that such register contrasts are tightly related to phonation differences in Shanghai Chinese (Cao and Maddieson 1992, Ren 1992, Zhu 1999). More specifically, L-register tones occur only on syllable with breathy voice (or murmur; I shall use them interchangeably), whereas H-register tones are compatible with syllables with modal (clear) phonation. The tone-phonation correlation is rooted in a well-known fact: breathy voice

---

[10] The recording was made by one female speaker PYL, aged 30 in 2001. The wordlist was given in Chinese characters and was read in isolation (Ping 2001b: 20-21). Those recorded monosyllables are open syllables, as far as I can tell. The acoustic measurements were conducted with KAY CSL 4300B.

lowers F_0 (Hombert 1978, Hombert et al. 1979, Gordon and Ladefoged 2001, Silverman 2002).

In addition, Ren (1992), through a series of acoustic measurements, perception tests, physiological investigations (including Rothenberg Mask Experiment and Fiberoptic, and Transillumination experiments), arrived at the conclusion in (18). In brief, he argues that such contrasts are not manifested either on the vowel or on the tone because murmur is most salient in the vocalic onset and fades away before the middle point of a vowel. In particular, if murmur were represented with a feature under the tonal node (Yip 1993), it would be expected that breathy voice lasts throughout the entire tone-bearing portion. Therefore, phonation is the onset consonant's inherent property (marked with ‘ ̤ ’ underneath a consonant). Zhu (1999) further notes that the modal vs. murmur distinction is attested in sonorants and onsetless vowels as well.

(18) Phonation contrasts among the obstruent categories (Ren 1992: 150)

	Initial Position		Non-initial position	
	Phonation type	Tone	Phonation type	Tone
paa	More adducted Voiceless	H-register	More adducted Voiceless	Neutralized
p̤aa	More abducted Voiceless	L-register	More abducted **Voiced**	Neutralized
p^haa	Most abducted Voiceless	H-register	Most abducted Voiceless	Neutralized

As (18) indicates, the phonation distinction is restricted to word-initial position. In non-initial position, murmur is lost, resulting in obstruent voicing and more importantly, *tonal neutralization*. Recall that this was dubbed “non-initial tone loss” in §1.2.4. As we will see in §2.3.4, this specific property plays a central role in understanding the motivation for SH tone sandhi and more generally, for Wu Chinese tone sandhi.

With these assumptions in mind, let us now turn to look at the normalized F_0 trajectories and phonetic lengths for each tone.

2.3.1.1 Acoustic properties of Shanghai Chinese citation tones

In this section, some essential phonetic properties of SH citation tones are described, in particular, F_0 and duration. All data are based on Ping's (2001b) experimental results. The normalized F_0 curves are plotted against normalized time and so are the logarithmic Z-score (LZ) normalized F_0 trajectories. LZ normalized values are the primary method to compare F_0 contours among sandhi and sandhi tones since SH also exhibits the positive skew in F_0 distribution (§1.3.1). As a notational convention throughout this thesis, H-register tone is represented with filled shape and L-register tone hollow shape.

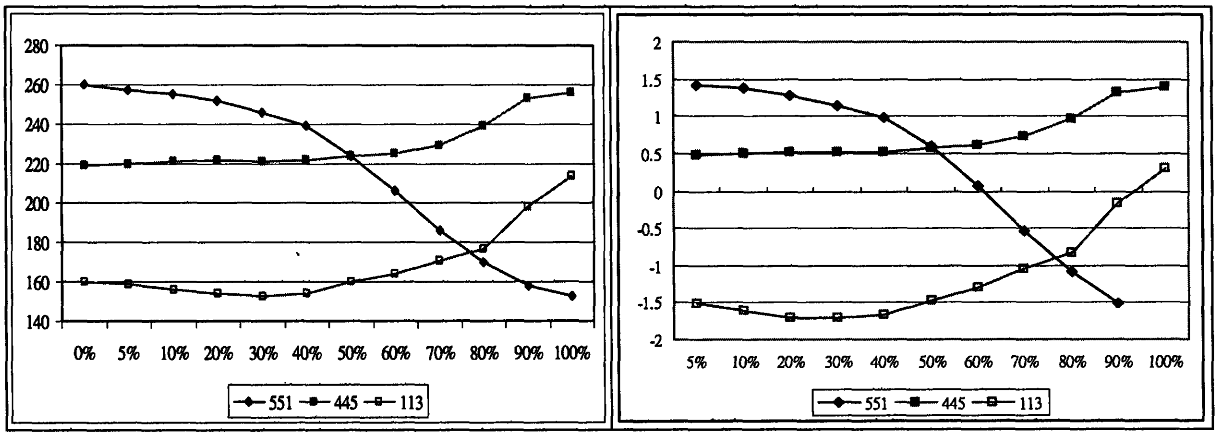


Figure 2-1 Normalized F_0 (in Hz) and LZ normalized F_0 for citation long tones in Shanghai Chinese

From the above F_0 tracings, it should be evident that the two rising tones are both of the delayed rise type (§1.3.2). For Tone 2 (455) and Tone 3 (113), the first half of the pitch contour is very flat. At the mid point, the F_0 curves begin to rise to reach the high endpoint. On the other hand, Tone 1 (551) looks like a delayed falling tone (§1.3.2). It can be seen from the above diagram that the slope from 0% to 50% of normalized duration is significantly shallower than that from 50% to 100%, as far as Tone 1 (551) is concerned.

The (LZ-)normalized F_0 tracks for the two checked tones in isolation are presented below. Tone 4 (55) is a high level checked tone. The final pitch jump is presumably due to the glottalization. Tone 5 (23) is a checked rising tone. As we can see, 23's contour shape is of the continuous rise type (§1.3.2). That is, no obvious F_0 plateau can be identified.

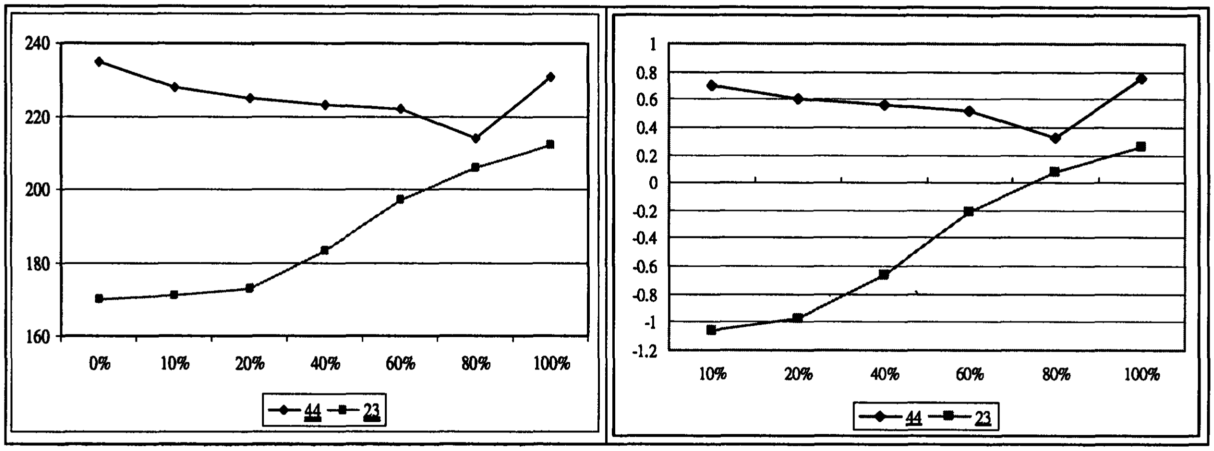


Figure 2-2 Normalized F_0 (in Hz) and LZ normalized F_0 for citation checked tones in Shanghai Chinese

The average rime duration for the citation tones is shown in the bar-plot in Figure 2-3. The error bar indicates one standard deviation. As is well-known, rising tones are cross-linguistically longer than falling tones and level tones (Ohala and Ewan 1973, Sundberg 1973, 1979, Ohala 1978, Xu and Sun 2002). This generalization also holds in SH. Among the long tones, tone 551 is the shortest, while the duration of tone 445 and tone 113 is basically the same. Checked rising tone 13 is longer than checked high tone 44.

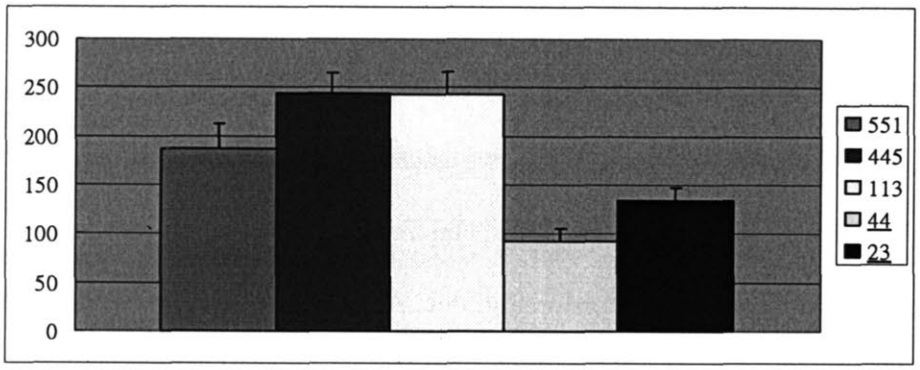


Figure 2-3 Duration of rime (in ms) for citation tones in Shanghai Chinese

Having described the basic acoustic properties of SH citation tones, we turn in the following section to defining the tone levels in SH.

2.3.1.2 The five-point scale in Shanghai Chinese

From the foregoing data section, the generalization is that SH tones can be represented on the basis of the following conversion table of LZ score to the five-point scale. The aim is to describe tone and tone sandhi in a more objective manner. As a first approximation, a 0.5 difference in LZ score (henceforth Δ_{LZ}) makes a pitch interval in SH. A raised exclamation mark '$^!$' denotes an extra low tone level, which is found only in sandhi contexts (cf. Patterns A and B Figure 2-4).

(19) Conversion of LZ score to the five-point scale in Shanghai Chinese

LZ-score	Five-point scale	Tone level
1 (or above)	5	H
0.5	4	M
0	3	M
-0.5	2	L
-1	1	L
(-1.5 or below	!1)	

Notice, however, the five-point scale does *not* mean that there are five contrasting tone levels. Rather, I posit that SH has three phonological tone levels, H, M, and L. This is mainly because the pitch shift size of Tone 2 (455) is around 1 in LZ, as shown in Figure 2-1, assuming that pitch excursion size is obtained between the 0% and 100% normalization points.[11] For the *citation* tone inventory, if a tone has a pitch excursion of 1 in LZ, we say that this tone is a contour tone.

Having discussed the phonetics and phonology of the SH citation tone inventory, we now proceed to the data section for disyllabic tone sandhi.

2.3.2 Disyllabic Tone Sandhi

This section provides a phonetic description of SH disyllabic tone sandhi. Like many other varieties of Wu Chinese, tone sandhi in polysyllabic compounds features initial/left-dominant or final/right-dominant tone sandhi. In SH, the initial-dominant sandhi can be used in words and phrases of any syntactic structure (hence was called "general sandhi rules" in the descriptive literature), while final-dominant is used for syntactic configurations such as Subject-Predicate and Verb-Complement (hence termed "restricted sandhi rules"). In this section, we will focus on

---

[11] In chapter 4, I will use different criteria to measure pitch excursion size in Shaoxing Chinese, also a Wu Chinese language. This is because the tone inventories in Shaoxing Chinese are much larger than those in SH. It will not make a crucial difference to use extreme normalization points for measurement in SH.

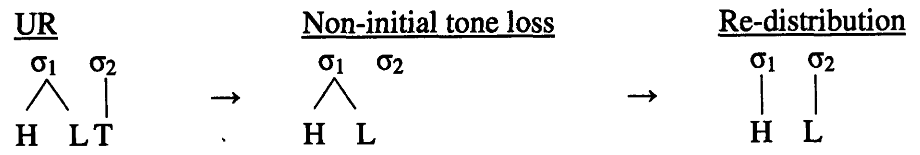
the disyllabic tone sandhi patterns and discussion and analysis of trisyllablic and quadrisyllablic tone sandhi will be postponed to §5.5.

The phonetic data reported here are of the initial-dominant sandhi. All of the recorded disyllabic words are Noun-Noun compounds or lexicalized Adjective-Noun phrases (in isolation). See Ping (2001b: 20-21) for the wordlist. After we go through the full range of tone sandhi data, this section will be closed by raising issues for the analytical section

2.3.2.1 Two types of tone re-distribution in Left-dominant sandhi

The initial-dominant tone sandhi in SH, or Initial Prominence, has been conventionally analyzed as deletion of non-initial tones (Duanmu 1990) or positional faithfulness in OT terms (Yip 2002, Li 2003). This is because in polysyllabic combinations, the overall tone patterns are determined by the initial tone. For the present purpose, let us use a theory-neutral term "non-initial tone loss" to describe this phenomenon (cf. tonal neutralization in (18)). Derivationally speaking, we see in (20) that re-distribution of the initial underlying tones takes place after the non-initial tone loss, yielding the general surface patterns in which the initial tonal contour superimposes on the entire disyllabic domain.

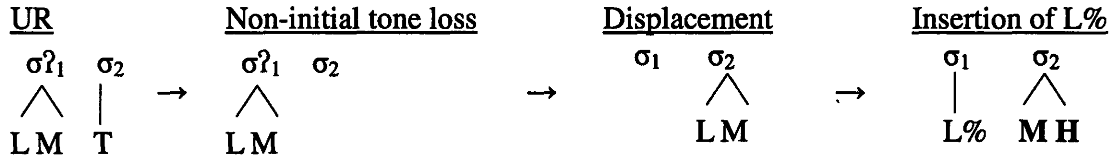
(20) General schema of Shanghai Chinese disyllabic tone sandhi



(T=any tone)

Yet another little-discussed type of the SH disyllabic tone sandhi is exemplified as follows. In this type, the checked rising tone (Tone 5: 13) moves *as a unit* to the following syllable and a boundary tone L% is inserted in the initial syllable. It is important to note that the L-register checked rising tone is $\widehat{\mathrm{LM}}$ (or $\underline{13}$) realized as an *H-register rising tone* $\widehat{\mathrm{MH}}$ (or $\underline{34}$) on the surface. Notice further that register-raising is not always attested in sandhi final position. For example, L tone in (20), belonging to L-register, occurs in this environment as well. See the corresponding F_0 tracings in Figure 2-4 (Patterns A and B) for validation.

(21) Contour displacement of the checked rising tone (Tone 5: $\underline{13}$)



(T=any tone; note also that the glottal stop coda drops in intersyllabic position.)

With this discussion in mind, the acoustic properties of Shanghai Chinese disyllabic tone sandhi are presented in the subsequent section.

2.3.2.2 Acoustic properties of disyllabic tone sandhi

Our discussion begins with Tone 1-initial patterns. Tone 1 is a high-falling tone, transcribed as 551. The table in (22) summarizes all of the attested two patterns. In the UR column (Underlying Representation), citation forms are given, followed by the SR column (Surface Representation). The initial tones and the final tones are separated by a hyphen. Mean duration of each tone is given and the standard deviation is parenthesized. Finally, N means number of tokens.

(22) Tone 1 (551) in initial position

	UR	SR	Rime duration in ms (SD)	N
A	551-Any Long Tone	55-1'1	147.7 (52.4)-136.4 (38.9)	26
B	551-Any Checked Tone	55-1'1	174.4 (24.8)-103.2 (24.7)	15

Regarding the F_0 trajectories, data are arranged in the following format. The normalized and LZ normalized on the left hand side denotes the initial tone, followed by a blank region. The blank region represents a hypothetical interlude consonant and does not stand for the duration of the actual interlude consonant. Finally, the F_0 curves on the right represent the second tone.

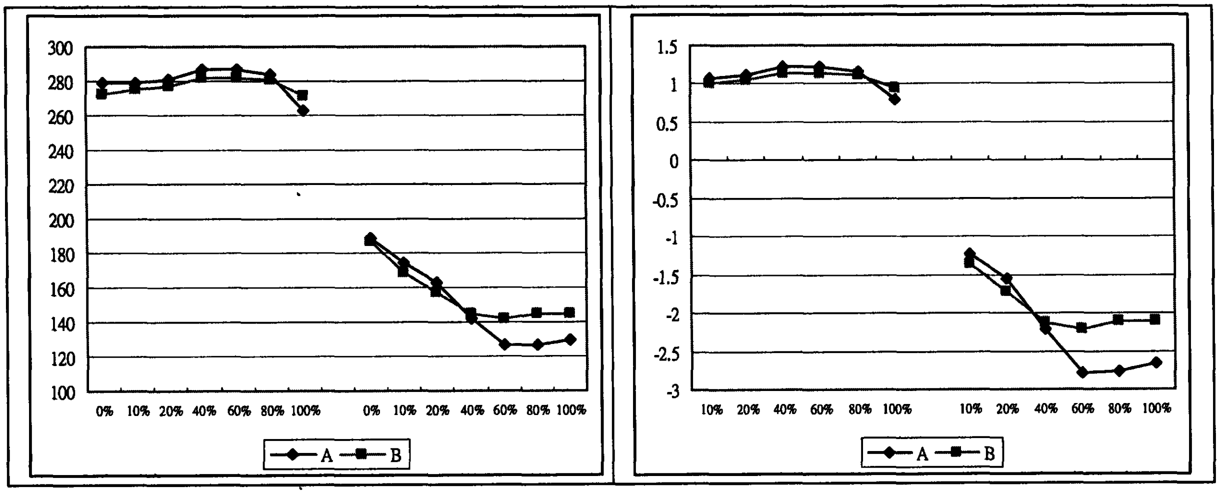


Figure 2-4 Normalized F_0 in Hz (left) and LZ normalized F_0 (right) for Patterns A and B

As we can see from the diagrams in Figure 2-4, 551 "splits" into two parts and the H peak stays on the initial syllable while the L part extends to the final syllable It should be noted that the L tone on the final syllable looks like a falling tone. But notice that the pitch range is extremely low at the end of the second syllable. Therefore, this can be treated as downgliding of the low tone, which is well-attested in tone languages (Maddieson 1978).

The case of the two long rising citation tones, 445 and 113, are presented below. Again, we see that the underlying rising contours on the initial syllable are redistributed over the entire disyllabic domain. More importantly, the final syllables are *longer* than the initial syllable, as evidenced in Patterns C and D. This fact contradicts the well-established generalization according to which the final tones are shorter than the initial tones in SH disyllables (e.g. Zee and Maddieson 1980, Duanmu 1993, Zhu 1999).

(23) Tone 2 (445) / Tone 3 (113) in initial position

	UR	SR	Rime duration (SD)	N
C	445-Any Long Tone	33-44	143.1 (55.2)-171.2(51.6)	45
D	445-Any Checked Tone	33-44	143.8 (34.8)-130.6 (23.3)	17
E	113-Any Long Tone	11-44	144.9 (34.2)-177.1 (38.5)	27
F	113-Any Checked Tone	11-44	160 (35.4)-130.3 (28.6)	17

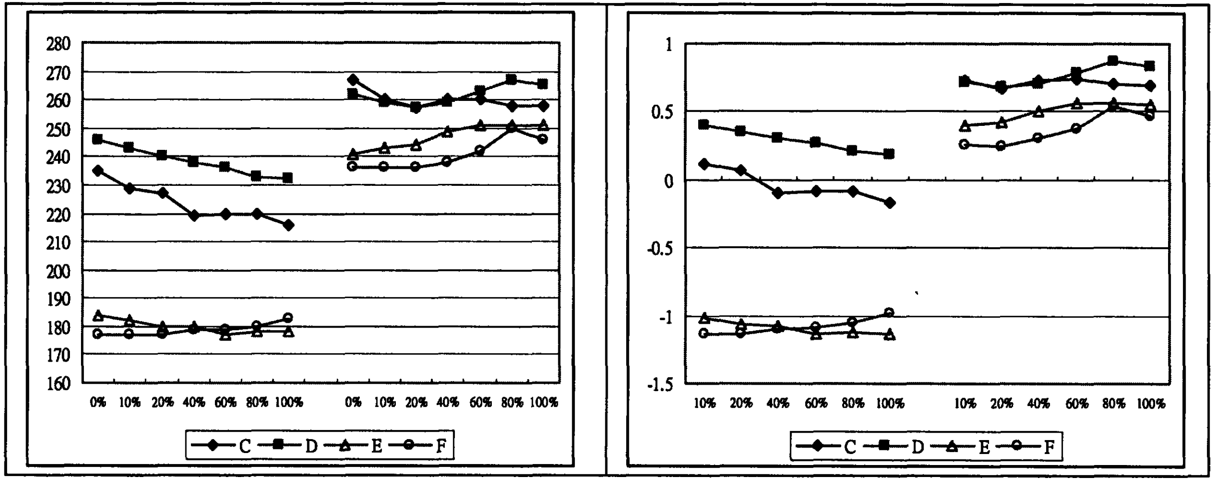


Figure 2-5 Normalized F_0 (left) and LZ normalized F_0 (right) for Patterns C, D, E, and F

Furthermore, one may wonder if it is appropriate to treat the seemingly falling contours of the initial tones for Patterns C and D as level tones. As we have discussed earlier, F_0 intervals are wider in the upper part of the pitch range due to the positive skew effect (§1.3.1). This view is further confirmed by the fact that the initial tonal contours are considerably flat in the lower range of the tonal space. I have posited in the preceding section that a 1 difference in LZ is the minimal pitch excursion size for a pitch level in the citation tone inventory. For sandhi tones, it appears that a 0.5 difference is more adequate. This proposal gains further support from the following data on the checked tone-initial patterns.

In Figure 2-6, Pattern I features the only attested contour tone in sandhi position, 34. Recall our discussion of contour displacement in (21). The underlying rising checked tone 13 moves to the following syllable, instead of splitting into the L and H parts, which is attested in all long tone-initial combinations (i.e. Patterns A-F).

Finally, the rise of sandhi tone 34 is about 0.5 in LZ, or approximately 30 Hz. Thus, I propose that 0.5 in LZ is taken as the minimally sufficient pitch excursion size for contour tones in SH.

(24) Checked tone in initial position

	UR	SR	Rime duration (SD)	N
G	Tone 4 (44)-Any Long Tone	33-44	77 (5.8)-199 (35.2)	27
H	Tone 4 (44)-Any Checked Tone	33-44	76.3 (18.5)-134.8 (34.5)	18
I	Tone 5 (23)-Any Long Tone	**11-34**	83.7 (17.5)-199.3 (35.8)	26
J	Tone 5 (23)-Any Checked Tone	11-44	70.4 (18.1)-120 (24.8)	18

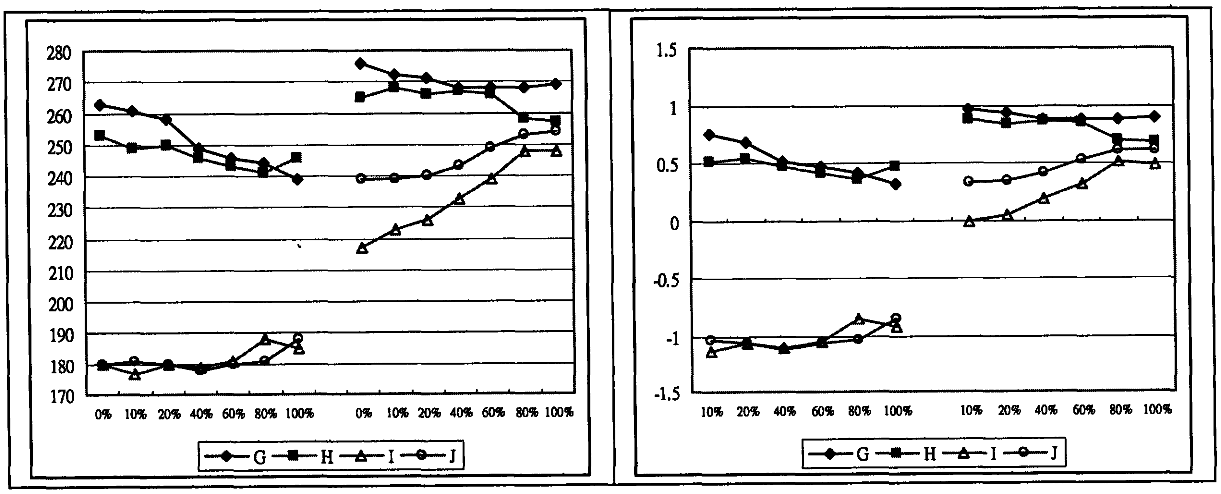


Figure 2-6 Normalized F_0 (left) and LZ normalized F_0 (right) for Patterns G, H, I and J

2.3.2.3 *Discussion*

To sum up, all 10 patterns in SH disyllabic tone sandhi are collocated below. Aside from the falling contour of the final L tone in Patterns A and B, which is caused by the well-known downgliding effect, it should be obvious that only tone 34 (i.e. the second tone in Pattern I) can be treated as a contour tone because this tone has the largest pitch excursion. As we have discussed above, the fluctuation of the other sandhi tones are all within the threshold value, 0.5 in LZ, or approximately 20~25 Hz.

More importantly, it is clear that the L-register checked tone 13 *raises to the H-register* in sandhi contexts, surfacing as 34. The F_0 contour begins at around 220 Hz, corresponding to 0 in LZ. It appears that there is no obvious motivation for this *register-raising* becasue L-register tones are *not* impossible in sandhi final position, as evidenced in Patterns A and B. This fact precludes an analysis appealing to insertion of register feature [+U]. As we will see, the core argument of contour correspondence hinges on this crucial fact.

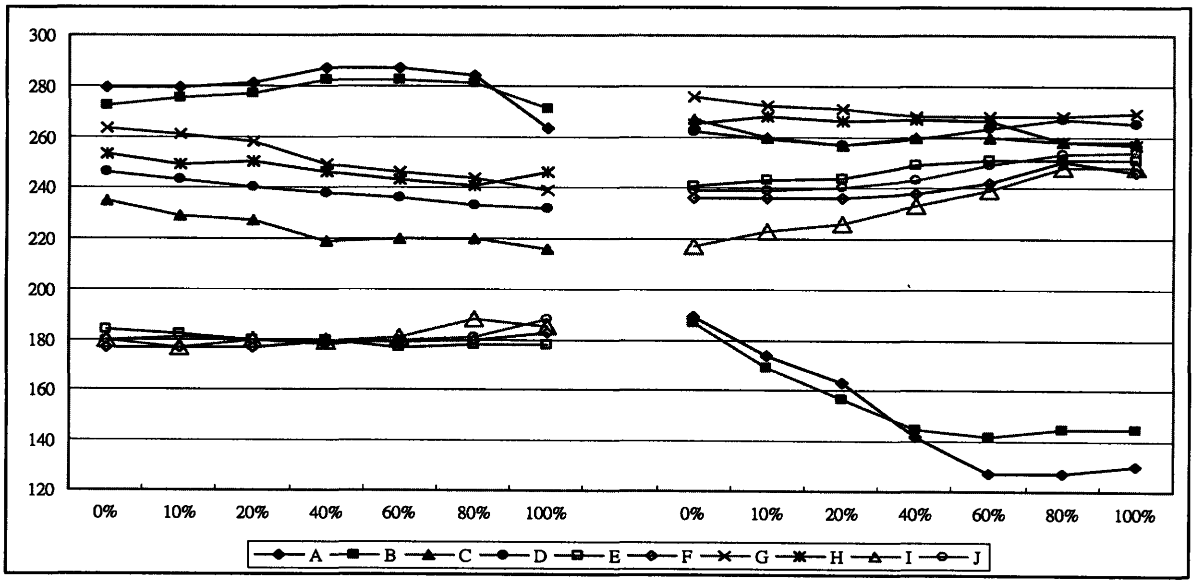


Figure 2-7 Normalized F_0 for all disyllabic tone sandhi patterns in Shanghai Chinese

Finally, the following two histograms present the clustered and stacked duration of i) the initial rime, ii) the C interlude, and iii) the final rime for the 10 disyllabic tone sandhi patterns in SH. As we can see in Figure 2-9, the canonical duration of a disyllable is approximately 400 ms (including the C interlude). Also, it appears that the initial syllables are not always provided with longer length, in particular, Patterns C and E.

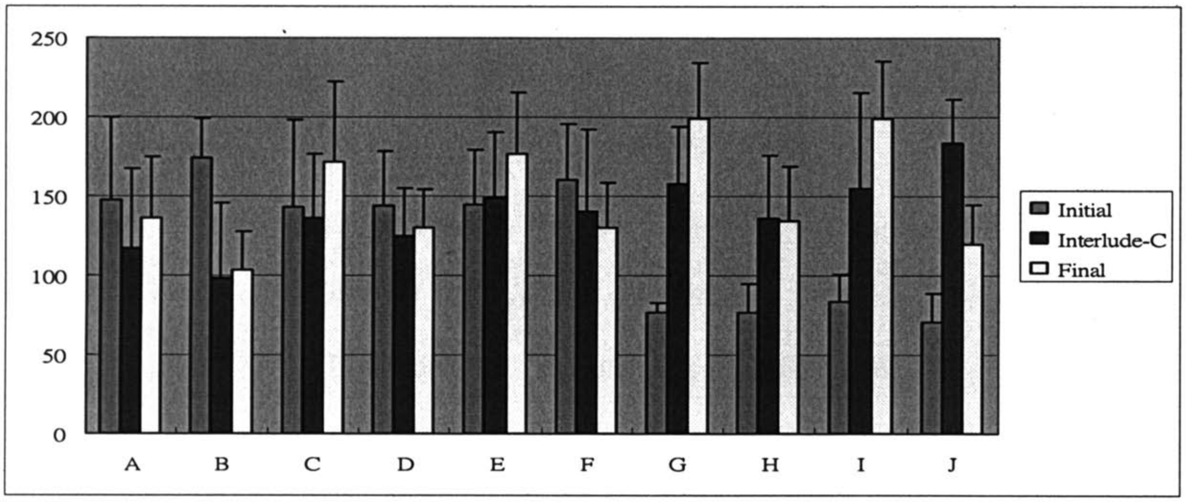


Figure 2-8 Clustered duration of rimes (in ms) for all disyllabic tone sandhi patterns in Shanghai Chinese

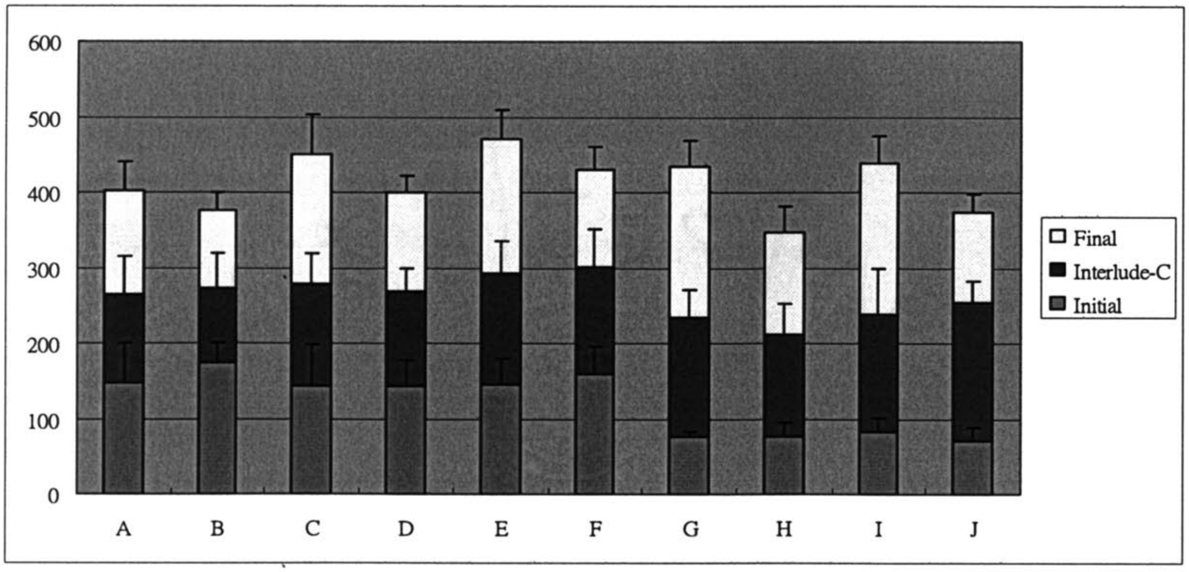


Figure 2-9 Stacked duration of rimes (in ms) for all disyllabic tone sandhi patterns in Shanghai Chinese

This completes our description and discussion of acoustic properties of SH disyllabic tone sandhi.

Bearing this range of data in mind, we now set out to analyze the central issue in this section: register-raising of the checked rising tone in terms of the familiar OT constraint families, in particular, unit correspondence constraints. A contour correspondence-based account and an analysis of the motivation for "non-initial tone loss" discussed in (20) and (21) will be separately provided in §2.3.4.

2.3.3 Against Unit Correspondence

This section is an attempt to account for register-raising of Tone 5 (13) in sandhi from the point view of element-based correspondence. As we have seen in the preceding section, the phenomenon in question involves contour displacement from the initial syllable (i.e. Tone 5's lexical source) to the final syllable (i.e. Tone 5's host on the surface), as indicated in boldface below. Tone 23 moves as a unit, while the initial syllable (its lexical source) is inserted with a low boundary tone, represented with '%' in (25). Note also that T means any *long* tone.

(25) Contour displacement and register raising (cf. Pattern I in Figure 2-6 and Figure 2-7)

a. **13**-T → 11$_{\%}$-**34**
b. [**-U, LH**]-T → [-U, L]$_{\%}$-[**+U, LH**]

In (25)b, I show that register-raising can be captured by insertion of [+U] in a multi-planar tonal node representation (Yip 1980, Bao 1990, 1999, among many others). As we have briefly discussed at the outset, however, "unrestricted" insertion of [+U] in sandhi final position yields the unattested pattern, as shown in (26). If the sandhi final tones were all H-registered, *55-33 would be the optimal output forms, contrary to fact.

(26) Final tones are not uniformly H-registered (cf. Patterns A&B in Figure 2-4 and Figure 2-7)

551-T → 55-1$^{!}$1 *55-**33**
[+U, HL]-T → [+U, H]-[-U, L] *[+U, H]-**[+U, L]**

As a first approximation, the fact that a low tone [-U, L] is attested in sandhi final position indicates that DEP-(+U) should be highly ranked and active in preventing the low tone from changing to a mid tone. So the generalization is that H and L-register tones are both allowed in sandhi final position. That being so, a puzzle arises: how and why is register-raising in (25) motivated? For present purposes, let us follow the conventional analysis of left-dominant tone sandhi provided in Yip (2002) (see Li 2003 for a similar approach). HEAD-MAX-T requires that input head tones are faithfully realized (but notice that register feature [±U] is not included). If HEAD-MAX-T outranks *T, the prediction is that only initial tones (i.e. head tones) survive. Thus candidate (a) is ruled out. The anti-contour tone constraint *CONTOUR eliminates candidates (b) and (c), ensuring that contour tones do not occur in polysyllabic compounds. Candidate (d) loses out because mid tone [+U, L] is not attested in sandhi final position. Ultimately, candidate (e) is chosen as the winning candidate.

(27) 551-445 → 55-1$^{!}$1

[+U, H_1L_2]-[+U, L_3H_4]	HEAD-MAX-T	*CONTOUR	*[+U]	*[-U]	*T
a. [+U, L_3]-[+U, H_4]	*!*		**		**
b. [+U, H_1L_2]-[-U, L_2]		*!	*	*	***
c. [+U, H_1L_2]-[+U, L_3H_4]		*!*	**		****
d. [+U, H_1]-[+U, L_2]			*!*		**
☞ e. [+U, H_1]-[-U, L_2]			*	*	**

According to the current ranking, sandhi output 34 [+U, LH] should be disallowed. The point is illustrated in the following tableau. Let us ignore for now why 34 [+U, LH] can overcome the anti-contour tone constraint *CONTOUR.

(28) $\underline{13}$-T → *$\underline{11}_{\%}$-13 (actual output: $\underline{11}_{\%}$-34)

$[-U, L_1H_2]$-T	HEAD-MAX-T	*[+U]	*[-U]
!☞ a. $[-U, L]_{\%}$-[-U, LH]			**
☹ b. $[-U, L]_{\%}$-[+U, LH]		*!	*

(where $[-U, L]_{\%}$ = the low boundary tone)

In order to motivate register-raising, it is conceivable that the most straightforward way is to propose a markedness constraint eliminating [-U, LH]. Since [-U, LH] is only allowed in citation and in initial sandhi position, the ranking HEAD-MAX-T » *[-U, LH] derives the desired results: an L-register rising tone is banned in sandhi non-initial position. There are, however, a handful of alternative possibilities to satisfy *[-U, LH]. For example, an L-register rising tone may undergo contour simplification to avoid violation of *[-U, LH], which is not the strategy adopted in SH. As we have seen, the rise of the L-register checked tone must be preserved on the surface. This can be achieved by the following constraint:

(29) MAX-CTU

> "Let CTU_y be the output correspondent of CTU_x. Tones associated with CTU_x in the input have correspondents associated with CTU_y in the output."

In addition to this, DEP-(+U) and IDENT-(-U) must be inert so as to motivate register-raising:

(30) Register-raising: [-U, LH] » IDENT-(+/-U)

[-U, LH]	*[-U, LH]	MAX-CTU	DEP-(+U)	DEP-(-U)	IDENT-(-U)
☞ a. [+U, LH]			*		*
b. [-U, LH]	*!				
c. [+U, L]		*!	*		*
d. [-U, L]		*!			

Moreover, Dep-(+U) must outrank Dep-(-U); otherwise wrong predictions will be made for the case below.

(31) 551-T → 55-1$^{!}$1 (cf. (27))

	[+U, HL]-T	Dep-(+U)	Dep-(-U)
	a. [+U, H]-[+U, L]	*!	
☞	b. [+U, H]-[-U, L]		*

The above analysis is problematic. We have to rank Ident-(+/-U) over *[-U, LH]. The ranking argument is that no L-register rising tone can survive if *[-U, LH] outranks Ident-(+/-U).

(32) Ident-(+/-U) » *[-U, LH], or no L-register rising tone can survive

	[-U, LH]	Ident-(-U)	*[-U, LH]	Max-CTU	Dep-(+U)	Dep-(-U)
☞	a. [-U, LH]		*			
	b. [+U, LH]	*!			*	
	c. [+U, L]	*!		*	*	

A ranking paradox emerges. If Ident-(-U) is active, it is surprising that register-raising is attested. The point is illustrated in (33). Let us again ignore for now how contour displacement is motivated.

(33) *[-U, LH]-in-Final

	13-113 [-U, L_1H_2]-[-U, L_3H_4]	*[-U, LH]-in-Final	Max-CTU	Ident-[-U]	*[-U, LH]	Dep-(+U)	Dep-(-U)
☞	a. $11_{\%}$-34 [-U, L]$_{\%}$-[+U, L_1H_2]			*		*	
	b. $11_{\%}$-13 [-U, L]$_{\%}$-[-U, L_1H_2]	*!			*		*
	c. 11-33 [-U, L_1]-[-U, H_2]		*!				*

(where L% = boundary tone.)

As seen, we are forced to stipulate a constraint against L-register rising tone in final position, *[-U, LH]-in-Final, and rank Max-CTU over *[-U, LH]. But *[-U, LH]-in-Final does not seem to solve the problem. L-register rising tone is not impossible in sandhi final position. Zee and Maddieson (1980) report that word-final LM$^{\uparrow}$ (= tone 5 (13)) is attested in SH trisyllabic tone sandhi. We know that arbitrary constraints are inevitable until we come up with a more viable solution. The above discussion shows that although at first sight register-raising can be neatly captured by insertion of [+U] in a straightforward fashion, the present data pose a non-trivial problem to unit correspondence. This is because L-register is the default register feature, through the established ranking: Dep-(+U) » Dep-(-U). Moreover, Ident-(-U) is active. Consequently, we cannot but resort to ad hoc markedness constraints. In sum, I would like to reiterate that Max-CTU is able to replicate the contour preservation effect only if insertion of a register feature is tolerated. Register-raising in SH contradicts an idea along this line because [+U] is *not* the default register feature, hence a non-trivial challenge to unit correspondence.

Before we launch into the analysis, we need to consider another approach in tone literature. We may follow the "Africanist" tradition and abandon the contour tone node (e.g. Duanmu 1994). I briefly sketched an analysis along this line in §1.2.3 (i.e. the "Invariance of Variation" phenomenon) and arrived at the conclusion that register-raising cannot be easily handled with this approach as well. Again, let us assume the ranking Head-Max-T » *T, requiring that only input head tones (i.e. the initial tones) survive on the surface. It turns out that the mapping L͡M → M͡H, or from input L-register rising tone to output H-register rising tone, violates *every* relevant unit correspondence constraints: Max, Dep, and Ident. Notice that L͡M must be dominated by Head-Max-T, or no L͡M would survive, which is contrary to the fact.

(34) Register-raising: $\underline{13}$-113 → $\underline{11}_{\%}$-34

	$\overgroup{\underline{L_1M_2}}$-$\overgroup{L_3M_4}$	Head-Max-T	*$\overgroup{LM}$	Max-T	Dep-T	Ident-T
!☞	a. $\underline{L}$%-$\overgroup{L_1M_2}$		*		*	*
☹	b. $\underline{L}$%-$\overgroup{M_3H}$	*!*		**	***	**

(where $\overgroup{LM}$=113/$\underline{23}$, $\overgroup{MH}$=34, %=boundary tone; checked tones are underlined.)

Even if we stipulate a constraint, *$\overgroup{LM}$-in-Final, the expected winner candidate (35)b still loses out because of the Dep violation. Notice also that being irrelevant, parenthesized L% is not considered here.

(35) Register-raising loses to contour simplification

	$\overgroup{\underline{L_1M_2}}$-T	*$\overgroup{LM}$-in-Final	Max-T	Dep-T	Ident-T
	a. ($\underline{L}$%-)$\overgroup{L_1M_2}$	*!			
☹	b. ($\underline{L}$%-)$\overgroup{M_2H}$		*	*	*
!☞	c. ($\underline{L}$%-)M_2		*		*
!☞	d. ($\underline{L}$%-)L_1		*		*

Demoting Dep-(H) does not solve the problem, either. This is because (35)c has the same violation profile with the expected winner (35)b, except that (35)c does *not* incurs a Dep violation. In other words, (35)b is always harmonically bound by (35)c.

It should now be clear that register-raising cannot be adequately accounted for in terms of unit correspondence and markedness constraints.

2.3.4 A Contour Correspondence Account

A closer look reveals the fact that initial syllables or syllables in isolation have a three-way tonal contrast, while tone levels are polarized in final position. More precisely, non-initial tones are generally either H (44) or L ($1^{1}1$). The only exception is 34 in Pattern I.

(36) Contextual variations of tonal contrasts (cf. Figure 2-7)

Citation/Sandhi initial	Sandhi final
5, 4/3, 2/1	4, 1 (**34** only attested in Pattern I)

I will offer an account shortly for the motivation why tonal contrasts are neutralized in sandhi final position. For the present purpose, let us exclusively focus on the analysis of register-raising. It is important to note that in sandhi final position pitch intervals 5 and 3 change to pitch interval 4, while pitch interval 1 remains 1.

(37)

Tone level	H	M	L	H	M	L
	1	2	3	1	2	3
Pitch level	5	3	1	4		1

As a first approximation, the reason why tone 5 (13) does not surface as 13, when moving to the sandhi final position is attributable to the fact that the pitch excursion size of 13 is too large for a durationally shorter syllable, which is about 100 ms shorter than the citation long syllables. As we have seen earlier, one pitch interval corresponds to 0.5 in LZ so 13 amounts to 1 in LZ, which is not likely in sandhi position: the maximal pitch excursion size is 0.5 in sandhi. In addition, the pitch shift size of the potential form 14 is too big for sandhi position, too. Taken together, I propose that register-raising is due to the interplay between RELCORR and tonal distinctiveness constraints. The tonal distinctiveness constraints proposed here primarily hinge on F_0, because it is well-established that this particular dimension serves as the main cue for tonal discrimination, Tonal distinctiveness is formulated with the following ingredients. First, under the assumption that Tone T has two temporal spans associated with a tone value, the tonal distinctiveness of contrasts should be maximized between a set of the spans on the same edge (i.e. either "both initial" or "both final"). Second, a specified minimal distance between contrasting targets is required, too.

(38) $\text{MINDIST}_{\text{TONE}} = \Delta_{\text{INITIAL}} > x \; \& \; \Delta_{\text{FINAL}} > x$ (abbr. $\text{MINDIST}_{\text{T}} = \text{I} \; \& \; \text{F} > x$)

> T_1 has two slices associated with a tone value, t_i and t_j, where t_i precedes t_j.
> T_2 has two slices associated with a tone value, t_p and t_q, where t_p precedes t_q.
> Assign a violation mark if $\neg(\Delta_{(ti\text{-}tp)} > x$ or $\Delta_{(tj\text{-}tq)} > x)$.
> where $x < 2$ in sandhi

If we rank the contour correspondence constraint for the "less than" relation, RELCORR(x<y), over the tonal distinctive constraint, $\text{MINDIST}_{\text{T}}$ I&F>2, then the prediction is that no contour flattening occurs. As we can see in candidates (c) and (d) below, 13 is neutralized with either 11 or 44, these output forms are penalized by RELCORR(x<y). Recall that level tones exhibit the "equal to" relation (x=y). A mapping, for example, from 23 (x<y) to 11 (x=y) incurs a violation of RELCORR(x<y). Candidate (a) loses out because the tonal space is polarized to 1 and 4 in sandhi position. The pitch intervals of 2 and 3 are positioned between 1 and 4, hence constitute insufficient distinctiveness with contrasting pitch intervals.

(39) Register-raising in sandhi position

11_1-13_2-44_3	RELCORR (x<y)	$\text{MINDIST}_{\text{T}}$ I&F>2
a. 11-23-44		*!*
☞ b. 11-34-44		*
c. $11_{1,2}$-44	*!	
d. 11-$44_{2,3}$	*!	

Let us now turn to look at the issue why only checked rising tone <u>13</u> undergoes contour displacement in sandhi. I propose that Contour tone can be "separable" by invoking IDENT-T-INITIALLONGSYLL. The line of reasoning is that non-checked syllables are longer than checked syllables. So it is reasonable to say that preservation of tonal identity is "more important" as far as non-checked long syllables are concerned. For the present purpose, we propose an all-purpose positional markedness constraint *InitialContour (or Yip's (2002) LicenceContour) because it is well-known that initial position is not privileged for contour tones (see Zhang (2002b) for argument against this approach). As we can see below, contour displacement occurs in response to *InitialContour and RelCorr (*x*<*y*).

(40) Contour displacement

$\underline{13}$-T	*InitialContour	Ident-(T) InitialLongV	RelCorr (*x*<*y*)	Ident-(T) InitialShortV
a. $\underline{11}$-44			*!	
☞ b. $\underline{11}_{\%}$-34				*
c. $\underline{13}$-$11_{\%}$				

By contrast, we see in the following tableau that Ident-(T)$_{\text{InitialLongV}}$ is active in preventing underlying contour tones on a long syllable from moving to the following syllable as a whole.

(41) Contour splits into two parts

113-T	Ident-(T)$_{\text{InitialLongV}}$	RelCorr (*x*<*y*)	Ident-(T)$_{\text{InitialShortV}}$
☞ a. 11-44	*	*	
b. $11_{\%}$-34	*!*		

Finally, I would like to address the following issue: what is the underlying motivation for "non-initial tone loss" in SH, or more generally, in Wu Chinese languages? Recall that breathy voice is neutralized in sandhi position. Cao and Maddieson (1992) has pointed out the correlation between non-initial tonal neutralization and loss of murmur in intersyllabic position. I argue that tonal neutralization can be captured in terms of multi-dimensional contrast preservation advanced in Flemming (1995, 2002). I first assume the following auditory dimension for phonation types, where the degree of adduction decreases from the leftmost creaky voice to the rightmost breathy voice.

(42)	Creaky	Modal	Aspiration	Breathy
	1	2	3	4

We conjoin the MINDIST constraints for the F_0 and voicing dimension. In (43), the voicing contrast is three-way in initial position, so the prediction is that distinctiveness of tone levels may not be that rigid, hence three-way contrast: H, M, and L, as in candidate (a) below.

(43) Three way contrast in initial position

	MINDIST = F0:1 & Voicing:2	MINDIST= F0:2 & Vocing:0	MAXIMIZE CONTRASTS
☞ a. H-M-L			√√
b. H-L			√
c. H-M	*!		√

On the other hand, we have learned that breathy voice is neutralized in intervocalic position. This fact indicates that tone levels are less in such a context if the relevant MINDIST constraints are properly ranked, as in candidate (b) below.

(44) Two way contrast in final position

	MINDIST = F0:2 & Voicing: 1	MINDIST = F0:1 & Voicing: 2	MAXIMIZE CONTRASTS
a. H-M-L	*!	*	√√
☞ b. H-L		*	√

2.3.5 Local Summary

In this section, I have demonstrated that contour correspondence and a scalar representation to tone sandhi provide a straightforward account for the register-raising puzzle attested in SH. The most surprising property of the phenomenon in question is that the contour is preserved regardless of violations of every relevant unit correspondence constraints, i.e. "Invariance of Variations." The present case study is the starting point of the overall project. Our plan for the analyses of various tone sandhi phenomena is outlined in the following section.

2.4 Discussion and Conclusion

In this chapter, I have laid out the general schema of contour correspondence and provided one case study concerning the register-raising in Shanghai Chinese disyllabic tone sandhi. The other three relations are left untouched. As for the 'greater than' relation, this relation will be extensively employed in the analysis of floating high tone metathesis in San Miguel el Grande Mixtec in §3.2. In chapter 4 on contouricity correspondence in Shaoxing Chinese is devoted to validate the "(non-)equal to" relation via a case study of the phonetics and phonology of tone sandhi in Shaoxing Chinese, a Northern Wu Chinese language.

So far our attention is restricted to preservation of the contour in the nucleus. As a further step, the next issue I would like to explore is the role of contour correspondence in wider contexts, in particular, contour preservation across a syllable boundary and a word boundary.

Furthermore, this also brings us back to a comparison of contour correspondence and Dilley's (2004) tone interval theory. As I have briefly mentioned at the outset, relational correspondence resembles tone interval in many respects. For example, one of the most important insights in Dilley (2004) is the novel way to represent syntagmatic relations between two tones (or "tone interval" in her formalism). She proposes that in tone interval, *specifications* of the phonological height relation include the following: "greater than," "smaller than," and "equals" (call it the "specification view of relation"). Consider a tone interval $[T_1\ T_2]$. If this tone interval is specified the 'greater than' relation, the surface realization of T_2 is higher than T_1. Relational correspondence is substantially different from her formulation in that relation is *not* a specification *per se*. Relation is established between two phonologically specified tonal slices. The following chapter aims to further tease apart relational correspondence and the tone interval theory by testing whether contour preservation is more rigid in the nucleus domain than across a syllable boundary. Under the specification view of relation, the height relations in (45)a-b are indistinguishable and are supposed to be protected by the same faithfulness constraint. The empirical facts nevertheless suggest the contrary. I will show that cross-linguistically there is a systematic asymmetry with respect to the rigidity of contour preservation in different prosodic contexts.

(45) Proximity between successive tones

a. [CV] b. [CV.CV]
/\ | |
HL H L (where [] = word boundary)

Chapter 3 Contextualized Contour correspondence

In chapter 2, I have laid out the general schema of contour correspondence in tone sandhi and analyzed an array of data in support of the existence of contour correspondence. The discussion was restricted to the (syllabic) nucleus domain. For a fuller picture, it is appealing to see whether contour correspondence is extendable to wider contexts. Therefore, the aim of this chapter is to investigate contour correspondence across syllable and word boundaries.

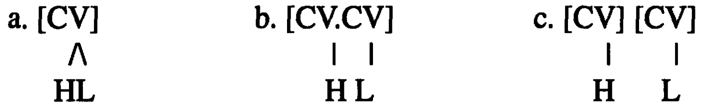
The organization of this chapter is as follows. In §3.1, I propose the intrinsic ranking of contextualized contour correspondence constraints. Two case studies are provided: the floating high tone metathesis in San Miguel el Grande Mixtec (§3.2) and the licensing condition on the creation of rising tone in Margi (§3.3). In the case study sections, I will show that a specific partial ranking of contour correspondence constraints plays a key role in accounting for the desired results. Comparisons with alternatives, in particular, Dilley (2004) and conclusions are presented in §3.4.

3.1 Contour correspondence Constraints and Their Intrinsic Ranking

3.1.1 Degree of Proximity Projected from Phonetics

In the previous chapter, our main concern was contour correspondence in the (syllabic) nucleus (or rime). The precedence relation between two temporal spans (associated with a tone value) is straightforward within a syllable. In wider contexts, however, precedence relations on the tonal tier can be established either *within* or *across* syllable(s), as graphically represented in (1)a and (1)b. Moreover, neighboring tones can also be separated by a word boundary in (1)c. The issue I would like to address is the following: does proximity between two successive tones play a role in tone sandhi?

(1) Proximity between successive tones



(Where square bracket = word boundary)

Importantly, the structures in (1) are indistinguishable if we consider the tonal tier alone, as is frequently assumed in autosegmental analysis. Contour correspondence makes very different predictions in this regard. The guiding intuition is that contour preservation should be more stringently enforced in a more *cohesive* domain. In other words, if H's and L's in (1) stand in relational correspondence (i.e. the 'greater than' relation), it is conceivable that change of this "greater than" relation is most resisted in the nucleus domain in (1)a and gives rise to the least repercussion across word boundary in (1)c. Of course, this idea is far from new, for example, Odden's (1994) proposal of adajaceny parameters. Suzuki (1998) proposes that proximity (i.e. the closer the elements are the stronger the interaction) and similarity (i.e. the more similar the elements are the stronger interaction) play important roles in the interaction of phonological elements. Archangeli and Pulleyblank (1994) also note that smaller domains enforce constraints more strongly than larger domains. Turning back to tone, we then want to ask: by what phonetic underpinnings is the relative stringency of contour preservation motivated in different proximities? It has long been noted that the fundamental frequency (F_0) is the major cue of tone perception (e.g. Gandour 1978). Gordon (2001) further points out that harmonics of a segment give good cues to the perception of the fundamental frequency. Vowels (on modal phonation) have a well-defined harmonic structure and sonorants have stronger acoustic manifestation of harmonics than obstruents. So F_0 is clearer in sonoarants (Gordon 2001, Zhang 2002 *inter alia*). Since the F_0 contour is generally uninterrupted throughout the vocalic portion of a nucleus, the most significant pitch information should be favorably encoded in this part. Conversely, the pitch contour across two syllables is tyoically interrupted by at least one consonant, be it sonorant or obstruent. It has been established that syllable onsets play no role in tone bearing ability (notice that F_0 lowering induced by the depressor consonants in Souther Bantu are not contrastive tones.). This is reflected in the following asymmetry: while coda sonorants may act as a tone-bearing unit, onset sonorants never do. The reason for this asymmetry is probably rooted in perception. House's (1990) psychoacoustic studies show that the hearer's sensitivity to pitch movement is significantly impeded in rapid spectral change, especially rapid increases in spectral energy such as that found in the syllable nucleus. In House's (1996) perception study of Thai tones, this proposal was revised and he claims that "tonal movement at syllable onset is perceptually coded differently from movement in the syllable nucleus or in the coda" in that "an F0 contour through the syllable onset or at the beginning of the nucleus is coded as a level tone while a contour through the nucleus or at the beginning of the coda is coded as a contour tone." It is plausible that onset is used for transition from preceding tone (neutral position for post-pausal syllables). Consequently, it is unlikely that salient pitch information is carried in the transition from the onset consonant to the vowel. (See Gordon 2001 and Zhang 2002b for more discussion and references cited therein).

Based on these findings, we are led to conclude that change of contour direction in the nucleus results in a more deviant percept, in comparison to change of contour direction across a syllable boundary (cf. the P-map approach in Steriade 2001). This is because pitch movement across an intervening consonant is not as robustly perceived as that on a sonorous rime. Confusability increases as the F_0 contour moves to a region of lower sonority. So a possible consequence of these perceptual effects on the phonological patterning is that contour preservation is *less* rigid across syllable boundary than in the nucleus. The present discussion is in line with the Stranded Tone Principle in §1.2.1. By way of a concrete example, consider the tonal adaptation patterns of medial epenthetic syllables in English-to-Yourba loanwrods (Kenstowicz 2006), where O=obstruent, R=sonorant.

(2) English loanwords into Yoruba

OR	HLL	OR	HHL
muffler	mófìlà	silver	sílífà

Kenstowicz observes that either the preceding or the following tone could be copied in a medial epenthetic vowel. The epenthetic vowel copies the tone of the vowel *flanking the sonorant.* In other words, the tonal contour across more sonorant segments is preferentially preserved in the loan forms, indicating that pitch information is more salient in a more sonorous span.

To encode the role of proximity/cohesiveness in phonology, this hypothesis is formalized by contextualizing the contour correspondence constraints and positing a fixed ranking for them. This ranking is projected from the speaker's knowledge according to which a contour change that is phonetically more salient is prohibited before a contour change that is less so. I propose three contextualized contour correspondence constraints, corresponding to the autosegmental representations in (1).

(3) The intrinsic ranking of the contour correspondence constraints (*to be defined* in §3.1.2)

RELCORR$_{\text{NUC}}$ » RELCORR$_{\text{WD}}$ » RELCORR

a. [CV]   b. [CV.CV]   c. [CV] [CV]
Λ | | | |
HL   H L   H   L   (Where square bracket = prosodic word)

RELCORR$_{\text{NUC}}$ is the *nucleus*-internal contour correspondence constraint, while RELCORR$_{\text{WD}}$ is the *word*-internal contour correspondence constraint. We have discussed earlier why RELCORR$_{\text{NUC}}$ always outranks RELCORR$_{\text{WD}}$. Not mentioned is RELCORR. This is the most generally defined Contour correspondence constraint, which is applicable in the case of contour preservation across word boundary. It should be reasonable to assume that RELCORR is lowest-ranked among the contour preservation constraints because boundary effects, e.g. pause, boundary tone etc., lead to more increase in confusability. I.e. pitch information is more obscured across a word boundary. In the following section, I set out to formally define contextualized contour correspondence constraints.

3.1.2 Defining Contextualized Contour correspondence

We have mentioned at the outset that within the (syllabic) nucleus domain, precedence relationship between tonal slices (or targets) is easily defined because the initial tonal target precedes the final one. Our discussion begins with the nucleus-internal contour correspondence, which is formulated below.

(4) RELCORR$_{NUC}$ (= nucleus-internal contour correspondence)

Let t_1 be a tone value contained within Rime R. Let S_1 be a temporal span associated with T_1
Let t_2 be a tone value contained within Rime R. Let S_2 be a temporal span associated with T_2.
S_1 precedes S_2.

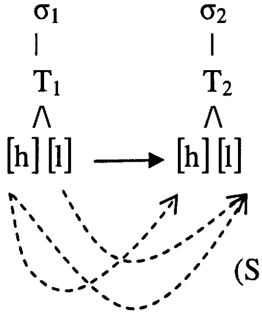
Let t_1' be the correspondent of t_1 in Rime R' and S_1', the temporal span associated with t_1' is the correspondent of S_1.
Let t_2' be the correspondent of t_2 in Rime R' and S_2', the temporal span associated with t_2' is the correspondent of S_2.
S_1' precedes S_2'.

Let $t_1 = x$, $t_2 = y$, $t_1' = a$, and $t_2' = b$.

v.	The 'greater than' relation:	If $x > y$, then $a > b$.
vi.	The 'less than' relation:	If $x < y$, then $a < b$.
vii.	The 'equal to' relation:	If $x = y$, then $a = b$.
viii.	The 'non-equal to' relation:	If $x \neq y$, then $a \neq b$.

In the previous chapter, we have postulated that the "elements" standing in contour correspondence are two hypothetical "temporal spans" or "slices" of the F_0 contour of Tone T. When two tones abut, however, given precedence relationship is more obscure. Consider now the following illustration of two successive falling tones.

(5) *Non-adjacent* tone slices do *not* stand in contour correspondence (=dotted arrow)



(Solid arrow = tonal slices standing in relational correspondence)

Given the four temporal spans in (5), we see that there are five logically possible precedence relations among *heterosyllabic* tones. I assume that *only immediately adjacent slices stand in* *contour correspondence*, i.e. the precedence relation marked with the *solid arrow* above is relevant for contour correspondence. More precisely, the right edge of the preceding tone and the left edge of the following tone. The reason is two fold. Recent studies on tonal coarticulation (1997, Xu 1994, 1995; Xu and Wang 2001) have shown that the adjacent pitch targets across syllables frequently interact but the pitch targets on the left edge of the preceding tone and on the right edge of the following tones do not. For example, given that there are two falling tones, H_1L_2-H_3L_4, tonal coarticulation is robustly attested between L_2 and H_3 whereas H_1 and H_3, H_1 and L_4, as well as L_2 and L_4 do not have significant interaction. These findings are in accord with the present assumption: only neighboring tonal slices across syllables stand in contour correspondence. Furthermore, it is well known that tones assimilate in much the same way as consonants and vowels do. But tones seem to spread only via autosegmental linking. To my knowledge, non-local tone spreading/assimilation seems unattested. For example, consider a tone sequence, T_1-T_2-T_3. I know of no documented case of 'action-at-a-distance' in tone sandhi where T_1 skips T_2, assimilating (or spreading) directly to T_3. Notice that long distance high tone displacement (or tone attraction) in West African languages is well-attested but one crucial property of this phenomenon is that low tone is phonologically inert. I.e. low tone is lack of specification so that high tone is able to do long-distance migration. (cf. Hansson's (2001) and Rose and Walker's (2004) proposal according to which consonant harmony is accomplished via segmental correspondence rather than autosegmental linking). Therefore, it should be appropriate to define contour preservation of tone in terms of *immediate* precedence.

Taken together, the "non-local" nature of tone suggests that the word-internal contour correspondence constraint be defined as follows. Suppose word ω has two tones T_1 and T_2 and ω' has two tones T_3 and T_4. ω' is the correspondent of ω.

(6) RELCORR$_{WD}$ (= Word-internal Contour correspondence)

Let t_1 be a tone value contained within Rime R_1. Let S_1 be a temporal span associated with T_1
Let t_2 be a tone value contained within Rime R_2. Let S_2 be a temporal span associated with T_2.

S_1 is on the right edge of R_1 and S_2 is on the left edge of R_2.

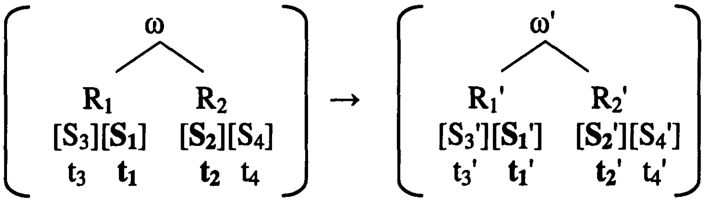
Let t_1' be the correspondent of t_1 in Rime R_1' and S_1', the temporal span associated with t_1' is the correspondent of S_1.
Let t_2' be the correspondent of t_2 in Rime R_2' and S_2', the temporal span associated with t_2' is the correspondent of S_2.

S_1' is on the right edge of R_1' and S_2' on the right edge of R_2'.

Let $R_1, R_2 \in$ Word ω, $R_1, R_2 \in$ Word ω'.
Word ω' is the correspondent of Word ω.

Let $t_1 = x$, $t_2 = y$, $t_1' = a$, and $t_2' = b$.

i.	The "greater than" relation:	If $x > y$, then $a > b$.
ii.	The "less than" relation:	If $x < y$, then $a < b$.
iii.	The "equal to" relation:	If $x = y$, then $a = b$.
iv.	The "non-equal to" relation:	If $x \neq y$, then $a \neq b$.



The most general contour preservation constraint, i.e. RELCORR, can be defined in a similar fashion, except that the slices which enter into a contour correspondence relation are the rightmost temporal span of the preceding word and the leftmost temporal span of the following word.

This completes our introduction to the contextualized Contour correspondence constraints. In the following sections, I set out to analyze empirical facts in support of contour correspondence. In particular, the upper part of the proposed intrinsic ranking in (3) will be established in the floating high tone metathesis in Mixtec (§3.2) and the lower part will be attested in the rising tone formation in Margi (§3.3).

3.2 Nucleus-internal Contour Preservation in Mixtec

As proposed in the preceding section, the central claim of this chapter is that all else being equal, contour correspondence is more stringently enforced in a more cohesive domain. This assumption was translated into the intrinsic ranking of the following contextualized contour correspondence constraints, which is repeated below.

(7) *RELCORR*$_{NUC}$ » *RELCORR*$_{WD}$ » RELCORR

This section deals with the upper part of the above ranking (in italics). The empirical evidence comes from the 'floating high tone metathesis' (after Goldsmith 1990) in San Miguel el Grande Mixtec. The point of interest is that this process is attested *only* in the word domain, rather than both in the (syllabic) nucleus and within the word. As a first approximation, we conjecture that this generalization is attributable to the very assumption: violation of contour correspondence is least tolerable in the nucleus since tone metathesis usually, if not always, leads to a change of contour direction. In other words, the reason why tone metathesis takes place word-internally, but not syllable-internally, is because contour preservation is less demanding in this environment. I will argue in the following sections that the partial ranking in (7) plays a crucial role in the floating high tone metathesis in Mixtec dialects.

3.2.1 Descriptive Preliminaries

Mixtec (or Mixteco, Mixtecan) is an Otomanguean language primarily spoken in the state of Oaxaca, Mexico. The Mixtec dialect under discussion is San Miguel el Grande Mixtec (henceforth SMG). In Mixtec, the notion of the *couplet* is central to understanding the canonical shapes of morphemes. Lexical words are viewed as couplets because of the fact that these morphemes conform to a highly restricted set of shapes: CVV and (C)VCV. According to Pike (1948) and Dyk and Stoudt (1965) , SMG has three tones, H, M and L. Contour tones do not occur. There are nevertheless 'long' vowels that look like a contour-toned syllable, e.g. k^wã̀ã́ 'yellow'. Note that 'long' vowels of this sort are treated as disyllabic in the Mixtec literature (Pike 1948, Macaulay 1996, but see Gerfen 1999: 21). For the present purpose, it should be safe to draw the conclusion that the 'one-tone-per-vowel' restriction applies across the board.

Consider now the nine (9) logically possible tonal patterns tabulated below. We find that tones can be freely concatenated with one another in a couplet except *LL (Pike 1948). I assume that the lack of the LL couplets is a lexical gap.

(8) Surface tone patterns in San Miguel el Grande Mixtec

HH	MH	LH
sáná 'turkey'	kūčí 'pig'	k^wã̀ã́ 'yellow'
HM	MM	LM
ñíʔī 'steam bath'	kēē 'go away'	mìnī 'puddle'
HL	ML	*LL
báʔù 'coyote'	kūtù 'nose'	

Of central interest is a phenomenon Pike termed 'tone perturbation.' This process is labeled as 'floating H metathesis' in Goldsmith (1990) and Goldsmith attributes it to the following language-specific property: certain couplets have a floating high tone suffix (represented with Ⓗ) in word-final position. Those patterns are listed in (9)a and their lexical representations are given in (9)b. Notice further that this floating high tone does not have any phonetic repercussion in isolation.

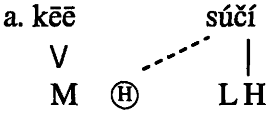
(9) Patterns with the word-final floating high tone

a. {HH Ⓗ} {HM Ⓗ} {MM Ⓗ} {ML Ⓗ} {LH Ⓗ}

b. MM with the floating high tone
kēē 'eat'
V
M Ⓗ

The existence of the floating high tone is evidenced in the minimal pair below. Note again that the verbs 'eat' and 'go away' are not distinguishable in isolation: *kēē*. As we can see, in (9)a, a final floating high tone docks on the initial syllable of the following morpheme, provided that no pause intervenes. By contrast, no tonal alternation occurs in (9)b. As Goldsmith (1990) suggests, it is appropriate to posit that the floating high tone triggers the alternation in (9)a (i.e. tone perturbation).

(10) Rightward migration of the floating high tone



'The child will eat.'

'The child will go away.'

As background knowledge, the brief description above should be sufficient. Before I present the full array of data, it is beneficial to confine our attention to the core data first. Below is a simplified illustration of the key distinction.

(11) The key distinction: CVV vs. CVCV

			Root type	*Would* contour shape change?
[T T Ⓗ]+[MH]	→	M**H**	CVV/CVCV	No, if metathesis takes place
[T T Ⓗ]+[ML]	→	**H**L	CVV	*Yes*, if metathesis takes place
[T T Ⓗ]+[ML]	→	M**H**	CVCV	*Yes*, if metathesis takes place

(Where T=any tone; square bracket=word boundary)

For the MH couplets, tone metathesis occurs on both CVV and CVCV roots. Notably, the overall contour shape is *not* changed among them: mid-to-high remains mid-to-high after the floating high tone docks on the second vowel. In contrast, for the ML couplets, *contour shape will be altered* if tone metathesis takes place. So it is important to note that CVV roots 'block' tone metathesis but the same process *is* attested on CVCV roots. This asymmetry indicates that *for CVCV couplets, contour shape change is tolerated but is NOT allowed for CVV couplets.* This disparity is captured by the central claim advanced at the outset:

(12) Preservation of tonal sequences is more rigid in the nucleus (=CVV) than in the word (=CVCV).

With this important notion in mind, the full array of data is presented in (13). Notice that T=any tone, except for those combinations without the floating high tone. Square brackets means word boundary. See (10)a and Pike (1948) for attested examples.

(13) Data on Ⓗ-association (see text for discussion)

a. Floating high tone replaces the first tone of the second word

UR		SR of the **second** word	Root type of the **second** word
[T T Ⓗ]+[HH]	→	$\underline{H}$H	All
[T T Ⓗ]+[HM]	→	$\underline{H}$M	All
[T T Ⓗ]+[HL]	→	$\underline{H}$L	All
[T T Ⓗ]+[LH]	→	$\underline{H}$H	All
[T T Ⓗ]+[LM]	→	$\underline{H}$M	All
[T T Ⓗ]+[MH]	→	$\underline{H}$H	CVʔV
[T T Ⓗ]+[MM]	→	$\underline{H}$M	CV$_i$V$_{i/j}$[1]/CVCV/CVʔV
[T T Ⓗ]+[ML]	→	$\underline{H}$L	CV$_i$V$_j$/CVʔV

b. Floating high tone metathesis

UR		SR of the **second** word	Root type of the **second** word
[T T Ⓗ]+[MH]	→	M$\underline{\mathbf{H}}$	CV$_i$V$_{i/j}$[2]/CVCV
[T T Ⓗ]+[ML]	→	M$\underline{\mathbf{H}}$	CV$_i$V$_j$/CVCV

For ease of discussion, I break down the description into two parts. First, if the second couplet begins with a high or low tone, the floating high tone appears on the initial mora. Second, if the second couplet begins with a mid tone, the surface tone patterns fall into two distinct types. For some root types, the floating tone docks on the initial syllable of the second word, which conforms to the patterns just described above (i.e. the last three items in (13)a). By contrast, for the other root types in (13)b, tone metathesis takes place: the floating high tone skips an initial mid-toned vowel and replaces a final vowel.

The next question is of course to understand the role of root type in tone metathesis. Recall that the couplets have two general shapes, CVV and (C)VCV. With respect to tone perturbation,

---

[1] The sandhi form HM optionally becomes HH on CV$_i$V$_i$ roots. See §3.2.2 for discussion.

[2] Tranel (1995, 1996) observes that no CV$_i$V$_i$ roots bear the MH pattern in Pike's data. Flipping through Dyk and Stoudt's (1965) SMG vocabulary, I found only one word with this pattern: **sāá** yācá 'crested **bird**.' I do not know if the word 'bird' is an exception. But it is true that words of this sort are extremely rare.

The constraint banning MH on CV$_i$V$_i$ roots plays a crucial role in Tranel's (1996) analysis. As we will see, our analysis does not need to have recourse to this specific constraint.

it is useful to distinguish two phonological classes of morphemes within each of the two categories, namely, i) words with contiguous vowels, i.e. CV_iV_i vs. CV_iV_j and ii) words with medial consonants, i.e. (C)VCV vs. CVʔ(C)V. Given this, it can be seen from the data set that tone metathesis systematically fails to occur on the glottalized roots CVʔ(C)V. This issue will be dealt with in §3.2.4.1. Second, there seems no consistency in the distribution of tone metathesis on $CV_iV_{i/j}$ and CVCV roots. In particular, CV_iV_j roots with ML do not pattern alike with CV_iV_i. See §3.2.4.2 for more discussion.

This completes the data section. We are now ready to launch into the analysis. The priority is given to the M-initial couplets, to which I turn in the following section.

3.2.2 M-initial Couplets

Tone metathesis occurs only when the initial vowel carries a mid tone. So it is tempting to posit that M is underspecified so that the floating high tone can skip a mid-toned initial vowel in certain couplets. I instead assume that H, M and L are all fully specified in the input. See §3.2.5 for more discussion on the underspecification account. We have learned that M is the only tone that is not replaced by the floating high tone. This provides a ranking argument for the following tonal faithfulness constraints.

(14) MAX-(M) » MAX-(H) // MAX-(L)

Furthermore, in order to minimize complications and concentrate on the phenomenon in question, let us simply assume that the floating high tone must be realized somewhere in the *following* word. Given these assumptions, let us now consider why floating high tone metathesis is blocked on CV_iV_i roots with ML (See §3.2.4.2 for discussion on CV_iV_j roots with ML).

(15) Floating H Metathesis is blocked on CV$_i$V$_i$ roots

CV$_i$ V$_i$ | | ⒽM L	RelCorr $(x>y)_{\text{Nuc}}$	Max- (M)	Max- (L)	RelCorr $(x>y)_{\text{Wd}}$
☞ a. CV$_i$ V$_i$ | | Ⓗ L		*		
b. CV$_i$ V$_i$ | | M Ⓗ	*!		*	*

Violation profile of the candidates w.r.t. RelCorr $(x>y)_{\text{Nuc}}$

Input	Output	Mapping	Violation
$(m_x l_y)$	$(h_a l_b)$	x>y → a>b	No
$(m_x l_y)$	$(m_a h_b)$	x>y → a<b	Yes

M and L constitute the 'greater than' relation on these CV$_i$V$_i$ roots. So the prediction is that contour change is prohibited because of the top-ranked RelCorr$(x>y)_{\text{Nuc}}$: In the nucleus domain (=CV$_i$V$_i$), the 'greater than' relation should be realized in the output. We see in candidate (b) that the floating high tone shifts to the second vowel, resulting in contour shape change: ML $(x>y)$ → MH $(a<b)$. So candidate (b) is eliminated because the input falling contour ML is not faithfully rendered in the output. In contrast, the floating high tone docks on the initial mora in candidate (a), yielding a high-to-low contour on the surface. Importantly, both the input ML and the output HL are of the 'greater than' relation.[3] Therefore, RelCorr$(x>y)_{\text{Nuc}}$ is satisfied for candidate (a). Consequently, the floating high tone chooses the initial mora as the landing site. We draw the conclusion that tone metathesis is 'blocked' in order to avoid violation of the nucleus-internal contour preservation constraint.

As for CVCV roots with ML, we have seen that tone metathesis occurs. The analysis is illustrated in the following tableau.

---

[3] The only difference lies in steepness. See chapter 5 for more discussion on the role of slope in tone sandhi.

(16) Floating H Metathesis takes place on CVCV roots

CVCV | | Ⓗ M L	RelCorr $(x>y)_{Nuc}$	Max- (M)	Max- (L)	RelCorr $(x>y)_{Wd}$
a. CVCV | | Ⓗ L		*!		
☞ b. CVCV | | M Ⓗ			*	*

Violation profile of the candidates w.r.t. RelCorr $(x>y)_{Wd}$

Input	Output	Mapping	Violation
$(m_x l_y)$	$(h_a l_b)$	x>y → a>b	No
$(m_x l_y)$	$(m_a h_b)$	x>y → a<b	Yes

Importantly, the nucleus-internal contour correspondence constraint (i.e. RelCorr$(x>y)_{Nuc}$) is vacuously satisfied here. Notice that the 'one-tone-per-vowel' principle has no exception in SMG. CVCV roots have two short vowels, so the nucleus-internal contour preservation is not active. By contrast, the word-internal contour correspondence constraint is violated in the winning candidate (b). In the input, the contour within the CVCV root is mid-to-low falling (i.e. ML) but the falling contour is not preserved in the actual output form. Since the word-internal contour correspondence constraint is dominated by Max-(M), it is predicted that the floating high tone skips the initial mid-toned vowel and replaces the second vowel. The floating high tone replaces the word-initial mid tone in candidate (a). This output incurs a fatal violation of Max-(M). We have mentioned that this ranking argument is based on the generalization that only the initial mid tone can be faithfully realized in some occasions of tone perturbation, suggesting that Max-(M) must outrank Max-(L). Therefore candidate (b) is the best way to avoid violation of Max-(M): the floating high tone shifts to the second vowel, replacing the second low tone, at cost of violation of lower ranked Max-(L).

From the above discussion, it should be clear that the floating high tone metathesis is conditioned by contour preservation. Tone metathesis takes place *only if* the top-ranked nucleus-internal contour correspondence constraint is satisfied. The absence of tone metathesis on CV$_i$V$_i$ roots is because the input ML should not change to MH. Contour change on a long vowel will result in a perceptually more deviant output form. So the speaker makes every effort to keep the contour shape. On the other hand, tampering with pitch contour interrupted by a consonant has less repercussion in perception. So the mid tone is preferentially retained on CVCV roots: ML → MH.

Regarding the MH pattern, our analysis predicts that the floating high tone always shifts to the second vowel because the nucleus-internal contour preservation constraint (i.e. RelCorr$_{\text{Nuc}}$) will not be violated in either CVV (=CV$_i$V$_i$ and CV$_i$V$_j$) or CVCV roots even though tone metathesis takes place. This is because the 'smaller than' relation (x<y) is still faithfully realized when the floating high tone is realized on the second vowel. The point is illustrated in the following tableaux.

(17) CVCV: Ⓗ MH → MH

<table>
<tr><th>CVCV
| |
Ⓗ M H</th><th>RelCorr
(x<y)$_{\text{Nuc}}$</th><th>Max-
(M)</th><th>Max-
(H)</th><th>RelCorr
(x<y)$_{\text{Wd}}$</th></tr>
<tr><td>☞ a. CVCV
| |
M Ⓗ</td><td></td><td></td><td>*</td><td></td></tr>
<tr><td>b. CVCV
| |
Ⓗ H</td><td></td><td>*!</td><td></td><td>*</td></tr>
</table>

(18) CVV: Ⓗ MH → MH

CVV | | Ⓗ MH	RelCorr $(x<y)_{\text{Nuc}}$	Max- (M)	Max- (H)	RelCorr $(x<y)_{\text{Wd}}$
☞ a. CVV | | MⒽ			*	
b. CVV | | ⒽH	*	*!		*

The analysis should be quite straightforward. The nucleus-internal contour shape is preserved because these couplets have a MH pattern. If the floating high tone docks onto the initial vowel, Max-(M) is violated. Consequently, tone metathesis is favored here.

Let us now turn to the last M-initial tone pattern: MM. This pattern is of the 'equal to' relation (x=y). From the data section, we have seen that tone metathesis does not occur in this pattern. The following tableau shows why this is the case. We need to invoke the anti-metathesis constraint Linearity to rule out candidate (a). This is because the 'equal to' relation will always be altered provided that the floating high tone must be realized in the second couplet. In addition to this, Max-(M) is unable to decide the winner, too. So the winner is the one that satisfies the lower ranked Linearity: candidate (a). Notice also that Linearity is ranked below Max-(M) since metathesis occurs in order to preserve the mid tone. The foregoing analyses still work because only RelCorr$_{\text{Nuc}}$ and Max-(M) play the key role in determining the winner for the ML and MH patterns.

(19) CVCV: Ⓗ MM → HM

<table>
<tr><th>CVCV
\/
Ⓗ M</th><th>RelCorr
$(x=y)_{Nuc}$</th><th>Max-
(M)</th><th>Linearity</th><th>RelCorr
$(x=y)_{Wd}$</th></tr>
<tr><td>a. CVCV
| |
M Ⓗ</td><td></td><td></td><td>*!</td><td>*</td></tr>
<tr><td>☞ b. CVCV
| |
Ⓗ M</td><td></td><td></td><td></td><td>*</td></tr>
</table>

Finally, Pike (1948) reports that there is an optional high tone spreading rule for the MM pattern on a long vowelled couplets (i.e. CV$_i$V$_i$ roots).

(20) Ⓗ kēē (MM) 'go away'→ kéē (HM) or kéé (HH)

Following the standard OT practice, I assume that this optional high tone spreading occurs because the two output forms above get tied in the ranking. To do this, I propose the following tonal identity constraint.

(21) Ident-Tone (abbr. Ident-(T))

> 'Let α be a tone-bearing unit in the input. Let β be the correspondent of α. The tonal specification of β is identical to the tonal specification of α.'

For the present case, I assume TBU is the mora (or vowel). As we can see below, Ident-T is violated twice in candidate (b) because the tonal specification of the two moras is not identical in the input and the output. As for candidate (a), the level contour is changed to falling, but this output incurs less violation of Ident-T. So optionality emerges from which of the two candidates get tied, as the following tableau illustrates.

(22) kēē (MM) 'go away'→ kéē (HM) or kéé (HH) in sandhi

CV V \/ ⒽM	RelCorr $(x=y)_{\text{Nuc}}$	Max-(M)	Linearity	RelCorr $(x=y)_{\text{Wd}}$	Ident-(T)
☞ a. CVV | | ⒽM	*			*	*
☞ b. CV V \/ Ⓗ		*			**

Finally, one may wonder about the behaviors of HH and LL on a long vowel. Recall that LL is a lexical gap in SMG and we do not expect to see any change on HH.

So far we have discussed the tone sandhi in M-initial couplets. I have proposed in §3.1.2.that RelCorr$_{\text{Nuc}}$ always outranks RelCorr$_{\text{Wd}}$. The ranking argument is based on the important observation that the floating high tone metathesis is strictly constrained by the requirement of the nucleus-internal contour preservation, while contour change is tolerated across a syllable boundary. As seen, this asymmetry is captured by ranking RelCorr$_{\text{Nuc}}$ over RelCorr$_{\text{Wd}}$.

The above is the most essential construct of our analysis of SMG floating tone metathesis. Regarding the H- and L-initial couplets, there are some complications even though tone perturbation is not attested in these patterns, to which I turn in the subsequent section.

3.2.3 H- and L-initial Couplets

This section is concerned with the tone sandhi data of H- and L-initial couplets. With regard to H-initial couplets, our observation is that no tone sandhi occurs. It is expected that the floating high tone will not change the input contour of H-initial couplets in any sensible way. The relevant analysis is not included.

The real challenge lies in tone perturbation in L-initial couplets. If the nucleus-internal "less than" relation must be kept in the output, candidate (a) below is wrongly selected as the winner.

(23) Wrong winner: Ⓗ LH → *LH

CVV | | Ⓗ LH	RELCORR $(x<y)_{\text{NUC}}$	MAX- (M)	LINEARITY	MAX- (L)
☹ a. CVV || ⒽH	*!			*
!☞ b. CVV || LⒽ			*	

Recall from the previous section that MH remains MH after tone perturbation. The present fact nevertheless suggests that rising contour (i.e. $x<y$) is more 'vulnerable' to contour shape change. In other words, RELCORR$(x<y)_{\text{NUC/WD}}$ should be demoted and be ranked below LINEARITY. This is not an unwelcome result, though. It has long been noted that rising tones are more effortful (Sundberg 1979, Ohala 1978, Xu and Sun 2002), requiring a longer duration to facilitate implementation. The inherent complexity of rising tones is also reflected in phonological patterning: the distribution of rising tone is cross-linguistically more restricted than that of falling tone (cf. two typological surveys of Gordon 2001 and Zhang 2002). Turning back to SMG, we may say that tone metathesis follows this trend to some extent. More importantly, the central claim of this chapter is not weakened. That a falling tone in the nucleus resists contour reduction or contour reversal does not necessarily entail that a rising tone in the nucleus has to do so. Likewise, the retention of a falling tone in the nucleus does not mean that the preservation of a rising contour across syllables must be less rigid. A more realistic assumption is that at least for the *same* relation, the nucleus-internal contour preservation should be more stringent than the word-internal contour preservation. For example, if the falling contour in the nucleus can be

altered, then the falling contour across syllables can also be changed, but *not* vice versa.

The revised analysis of the problematic case 'Ⓗ LH → HH' is provided below. As we can see, Linearity now prevents the floating high tone metathesis. Consequently, the floating high tone is realized on the initial mora. For comparison, I also include the 'Ⓗ MH → MH' pattern in the lower tableau. It turns out that Max-(M) triggers tone perturbation in this case.

(24) Ⓗ LH → HH / Ⓗ MH → MH

CVV | | Ⓗ LH	Max- (M)	Linearity	RelCorr $(x<y)_{\text{Nuc}}$	Max- (L)
☞ a. CVV | | ⒽH			*	*
b. CVV | | LⒽ		*!		

CVV | | Ⓗ MH	Max- (M)	Linearity	RelCorr $(x<y)_{\text{Nuc}}$	Max- (H)
c. CVV | | ⒽH	*!		*	
☞ d. CVV | | MⒽ		*		*

Up to this point, we have discussed and analyzed most of the SMG tone metathesis data. There is no denying that contour preservation is just one of the driving forces of tone perturbation. For example, Max-(M) also plays a substantial role in this regard. However, it is important to note that the CVV and CVCV distinction is especially difficult to analyze without recourse to contour correspondence: Ⓗ ML → **HL** on CVV roots but Ⓗ ML → **MH** on CVCV roots. This is again because the neighboring elements are not taken into account in unit correspondence. In addition to this, degree of proximity is not encoded in autosegmental representation. The SMG data necessitate both contour correspondence and proximity in tone sandhi. In other words, contour correspondence constraints must be contextualized and ranked in a fixed manner.

In the following section, for thoroughness I would like to address additional issues found in some minor types of tone perturbation.

3.2.4 Other Types of Couplets

3.2.4.1 Glottalized Roots

Not mentioned previously are CVʔV and CVʔCV roots (i.e. the glottalized roots). Interestingly enough, the floating high tone metathesis is *blocked* in this root type. This particular fact does challenge our analysis because the RELCORR$_{\text{NUC}}$ constraint prohibits tone metathesis only on CVV roots. By contrast, tone perturbation is allowed on CVCV roots because the nucleus-internal contour preservation is not relevant. Therefore, it is striking to see that metathesis *is* attested on CVʔ(C)V.

(25) The asymmetry of CVCV and CVʔ(C)V roots in tone perturbation

Goldsmith (1990) treats glottalization as associated to a consonant position on the skeleton (26)a. Hinton *et al.* (1991) argues for the vocalic approach (26)b. In other words, glottalization associates with the vowel. Finally, Macaulay and Salmons (1995) and Gerfern (1999) argue that glottalization is better treated as a floating feature [+constricted glottis] (26)c.

(26) Representation of Glottalization in Mixtec

a. Goldsmith (1990)   b. Hinton et al. (1991)   c. Macaulay&Salmons (1995)
CVCV   CVV   σ ʔ σ
||| Λ\   Λ/ \
TʔT   Tʔ T   CV V

It is generally agreed among previous researchers that glottalization interacts with tone in some fashion (Goldsmith 1990, Tranel 1995, 1996, Yip 2002). Some models of tonal representation treat tone/register features on a par with laryngeal features (e.g. Halle and Stevens 1971, Bao 1990, Duanmu 2000). So it is reasonable to assume that *postvocalic* glottal stops carry some tone feature (call it T).[4] Given this assumption, the glottalized/non-glottalized roots have different representations on the tonal tier.

(27) (Non-)glottalized roots

a. CVʔV   b. CVʔCV   c. CVCV
||| || | | |
MTL   MT L   M L

It appears that two association lines will be crossed if the floating high tone shifts to the second vowel. It is possible for the floating high tone to cross one tone (27)c, but crossing two linked tones is simply out, e.g. (27)a-b. This generalization can be captured by employing

---

[4] Hombert (1978) and many others have shown that postvocalic laryngeals do have an effect on pitch perturbation. Conversely, *prevocalic* laryngeals are not reported to affect pitch perturbation to a (phonologically) significant extent.

In Mixtec, word-initial glottal stops do exist. But they seem not to have any interaction with tones, that is, tones can cross them without any problem. As a tentative conclusion to this end, it should be safe not to represent prevocalic laryngeals on the tonal tier. The asymmetry regarding pre- vs. post-vocalic laryngeals in this respect definitely merits more study.

self-conjunction of LINEARITY (Smolensky 1995): LINEARITY&LINEARITY. This constraint is violated if and only if the floating high tone crosses two tones (or two association lines). So if this newly introduced constraint is undominated, then the prediction is that tone metathesis is blocked on glottalized roots. See below.

(28) Self-conjoined LINEARITY

CVʔV ||| ⒽMT L	LINEARITY& LINEARITY	RELCORR $(x>y)_{\mathrm{NUC}}$	MAX-(M)	LINEARITY	RELCORR $(x>y)_{\mathrm{WD}}$
☞ a. CVʔV ||| ⒽTL			*		
b. CVʔV ||| M TⒽ	*!			*	

Finally, an alert reader may wonder what might happen to a trisyllabic word. As a reminder, most of the lexical words in Mixtec are bimoraic (or disyllabic). So it remains to be seen if our analysis will make correct predictions if trisyllabic (monomorphemic) words exist at all: the floating high tone is supposed to dock onto the second vowel, not onto the third vowel.

3.2.4.2 CV_iV_j as Disyllabic Roots

The last tone sandhi pattern I would like to address is CV_iV_j roots with ML. Tone metathesis is attested in these roots. This is again unexpected if we regard CV_iV_i and CV_iV_j as like patterns (because both of them have two vowels in a row). It appears that RELCORR $(x>y)_{\mathrm{NUC}}$ is violated in the sandhi outputs of CV_iV_j roots.

(29) “Unexpected” tone metathesis in $\mathbf{CV_iV_j}$

Ⓗ žāù (ML) → žāú (MH) ‘cave’ (not *HL)

Hollenbach (2001) points out that from a comparative dialectological perspective, CV_iV_j roots in SMG were derived from CVCV roots, resulting from the diachronic process whereby intervocalic consonants undergo lenition and then drop. Gerfen (1999) also reports a parallel development in Coatzospan Mixtec. Furthermore, I found in Dyk and Stoudt’s (1965) SMG lexicon list that CV_iV_j roots consist of vocalic sequences such as [éà], indicating that the vocalic sequences in CV_iV_j roots are unlikely to be diphthongs (at least for some of them).

Therefore, it should be safe to conclude that couplets of this sort are better analyzed as disyllabic roots. For example, žāwù (but not žāù) ‘cave’ would be a less confusing transcription. Alternatively, it might well be the case that tone metathesis has been phonologized on CV_iV_j roots. More research is in need to clarify this issue.

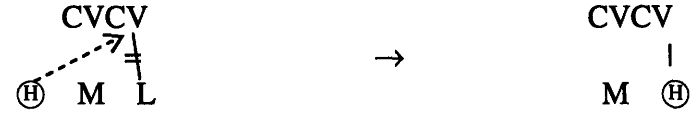
Before we close the discussion of the SMG data, it is necessary to see how floating high tone metathesis was analyzed in the former literature. As I have briefly mentioned earlier, it is tempting to analyze the peculiarity of M-initial couplets with underspecification. In the following section, I would like to review an early OT analysis proposed in Tranel (1996) and show that underspecification does not really solve the phenomenon in question.

3.2.5 Review of the Underspecification Account

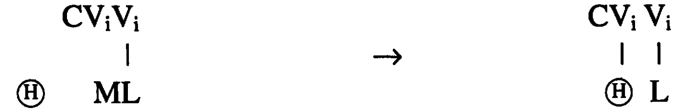
Tranel (1995, 1996) argues that floating high tone metathesis is due to the unmarkedness of the mid tone (Pulleyblank 1986, *inter alia*). That is, M is simply not present or is unassociated to a vowel in the underlying representation. The basic idea is that tone metathesis takes place simply because the mid tone is "transparent." The "transparency" effects of the mid-toned vowels can be stipulated through underspecification. Nevertheless, it appears that underspecification alone is not sufficient. The puzzle is that, other things being equal, in one context M is "transparent" (30)a and in yet another context M is *not* "transparent" (30)b.

(30) Transparent M vs. Non-transparent M

a. Transparent M



b. Non-transparent M



The analysis advanced in Tranel (1996) relies on an alignment constraint: TONE-LEFT, "*defining the floating tone as prefixes by demanding that they be on the left edge of their host's tonal tier.*" According to Tranel's assumption, both of the surface forms in (30) satisfy TONE-LEFT, because "the first vowel is toneless" in (30)a.[5] I assume that by "toneless," he means that the pitch value of the mid tone is determined by the interpolation between neighboring tones (see also Daly and Hyman 2005 for a similar proposal for the mid tone in Pioñle Mixtec).

Another important constraint in his analysis is the TONEPROMINENCEFAITH constraint (abbr. TONEPROMFAITH), which demands that "*an underlying toneless vowel must remain toneless in*

---

[5] Another plausible interpretation is that the mid tone simply has a 'tone target,' A tone target may have no underlying tone value specification and its surface value may be contextually conditioned or a default pitch value, normally the middle-range pitch. But under this assumption surface M tone is still subject to TONE-LEFT. See also Chen and Xu's (2006) proposal regarding the neutral tone in Mandarin.

the output." Tranel's (1996) analysis then proceeds as follows. In (31)a, the M-toned vowel is assumed to be toneless. Therefore, ToneLeft is not relevant even though the floating high tone replaces the low tone on the second vowel. Regarding (31)b, if the floating high tone docks onto the initial vowel, TonePromFaith is violated because an underlying toneless vowel must remain toneless in the output (presumably Dep-Tone and Ident-Tone rank very high).

(31) Floating high tone metathesis takes place on CVCV roots

CVCV | Ⓗ L	Tone-Left	TonePromFaith
☞ a. CVCV | Ⓗ (=MH)		
b. CVCV | | Ⓗ L		*!

(Notice that toneless vowel = M according to Tranel's assumption)

Furthermore, Tranel observes that CV_iV_i roots do not have the MH pattern,[6] so we may posit that the language-specific phonotactic constraint: *MH-on-CV_iV_i 'No MH pattern on CV_iV_i roots' is undominated. As we can see, this explains why the M is not "transparent" in (30)b or candidate (b) below.

---

[6] Tranel (1995, 1996) claims that CV_iV_i couplets with ML should be a lexical gap, at least according to Pike's data. In fn. 2, I have mentioned that only one example is found in Dyk and Stoudt's (1965) lexicon list, indicating that CV_iV_i couplets with MH are indeed extremely rare in SMG.

In Chalcatongo Mixtec, the distribution of floating high ton metathesis is basically the same as that of SMG. According to Macaulay (1996), MH is attested on CV_iV_i couplets and tone perturbation is also *blocked* on CV_iV_i roots with the ML pattern: Ⓗ MH → $\underline{\text{H}}$L. Therefore, the specific constraint against MH on CV_iV_i roots does not work in Chalcatongo Mixtec. By contrast, the analysis couched in contour correspondence can be extended to Chalcatongo Mixtec without any problem.

(32) Floating high tone metathesis is blocked on CVV roots

CVV | Ⓗ L	*MH-ON-CV$_i$V$_i$	TONE-LEFT	TONEPROMFAITH
a. CVV | Ⓗ (=MH)	*!		
☞ b. CVV | | ⒽL			*

(Notice that toneless vowel = M according to Tranel's assumption)

However, if the tone sandhi output of the MM pattern is taken into consideration, Tranel's analysis runs into problems. Consider the data and the tableau below.

(33) Ⓗ MM → HM or MH?

CVCV → CVCV or CVCV
| |
Ⓗ Ⓗ Ⓗ

CVCV Ⓗ	TONE-LEFT	Ⓗ-ON-INITIALV	TONEPROMFAITH
☞ a. CVCV | Ⓗ (=MH)		*!	*
b. CVCV | Ⓗ (=HM)			*

(Note that unspecified tone = mid tone in Tranel's analysis.)

As discussed earlier, TONE-LEFT should be satisfied with both of the candidates above because M is regarded as 'toneless,' hence is not subject to this alignment constraint. Notice further that LINEARITY (i.e. 'No Metathesis') must be irrelevant here because under his assumption there is no surface specification of M. Note that the actual output is HM, *not* *MH but the present ranking of constraints cannot decide the winner. So it is necessary to force the floating high tone to parse on the initial vowel of the word. To do this, we may stipulate a constraint: Ⓗ-ON-INITIALV. This new constraint requires that the floating high tone must dock onto the word-initial vowel. As seen, the current ranking successfully yields the expected output form. Turning back to CVCV roots with ML in (31) (=(34) below), however, the analysis makes an incorrect prediction: the floating high tone metathesis is supposed to be blocked because Ⓗ-ON-INITIALV outranks TONEPROMFAITH. Recall that TONEPROMFAITH requires that an underlying toneless vowel remains toneless in the output. In (34)a, the initial vowel is unspecified with tone, so tone metathesis is motivated because the floating tone avoids to dock on a toneless vowel. Now, under the pressure from the active constraint Ⓗ-ON-INITIALV, (34)b is wrongly selected as the optimal output.

(34) Wrong winner: Ⓗ ML → MH on CVCV roots

CVCV | Ⓗ L	TONE-LEFT	Ⓗ-ON-INITIALV	TONEPROMFAITH
☹ a. CVCV | Ⓗ (=MH)		*!	
!☞ b. CVCV | | Ⓗ L			*

(Note that unspecified tone = mid tone in Tranel's analysis.)

In summary, the underspecification account fails to capture the fact that only M is *not* "replaceable" in SMG. More importantly, tone metathesis is not due to the assumption according to which M is "transparent." As extensively discussed in the preceding subsections, tone metathesis is strictly constrained by contour preservation.

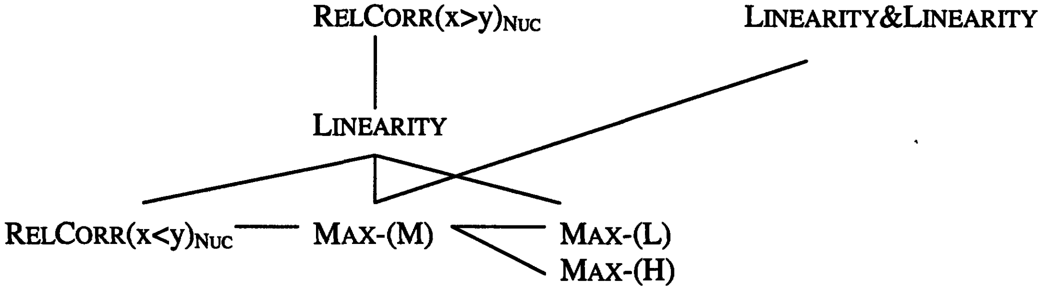
From a cross-linguistic perspective, SMG challenges the tonal prominence scale proposed in de Lacy (2002a): H ≻ M ≻ L as well as Pulleyblank's (2005) harmonic scale: M ≻ L ≻ H, where M is the most harmonic (least marked) tone, H is the most marked tone and L is intermediate. Either way, if these scales are converted into ranked faithfulness conditions, we *never* obtain: MAX-(M) » MAX-(H) » MAX-(L). Without this ranking, however, patterns of tone metathesis in SMG are incorrectly derived. We may simply postulate that languages differ in the tonal scales that they employ (e.g. Hume and Tserdanelis 2003), or await further research for a more viable soulation.The important point is that SMG is cross-linguistically unusual in that the preferential retention of the mid tone is attested if our analysis is on the right track.

3.2.6 Summary of this Section

In this section, I have argued that contour preservation must be considered in the analysis of floating tone metathesis in SMG. Furthermore, I have demonstrated that the upper part of the ranking proposed in (3): RELCORR$_{NUC}$ » RELCORR$_{WD}$ gains empirical support from this case study. Finally, I have also shown that an underspecification account makes wrong predictions. To this end, I conclude that an adequate analysis of SMG tone metathesis must refer to contour correspondence. In particular, the CVV and CVCV distinction must be recognized as the nucleus- and word-internal contour preservation.

In sum, the SMG data can be derived by the collated ranking of crucial constraints we have established so far.

(35) Collated ranking of crucial constraints



In the next section, let us move on to see how the lower part of the ranking helps us analyze the licensing condition on rising tone in Margi.

3.3 Word-internal Contour Preservation in Margi

The bulk of this section focuses on the lower part of proposed fixed ranking among contour correspondence constraints.

(36) RELCORR$_{\text{NUC}}$ » *RELCORR*$_{WD}$ » *RELCORR*

The prediction of the partial ranking under discussion is that contour preservation is attested within the word but is not imposed across words. The evidence in support of this claim comes from Margi: rising tones are created when glide formation occurs within a word, while rising tones are disallowed even though the same process, i.e. glide formation, takes place across words. This asymmetric condition on rising tone formation is difficult to analyze in terms of markedness because anti-contour tone constraints such as *RISE are not supposed to be voided only within the word domain. Instead, rising tone formation is better analyzed in terms of faithfulness. This is the starting point of our analysis.

3.3.1 The Stranded Tone Principle

Margi is a Chadic language of Northern Nigeria (Hoffmann 1963). On the surface, a syllable will bear one of four identifiable tones, High, Low, Rising, or Falling. Falling tones are "extremely rare" (Hoffmann 1963: 33), basically occurring in "expressive items" (Tranel 1992/1994: 115). In contrast, rising tones are robustly attested in this language. There are two sources for rising tone on the surface. One is lexically specified, i.e. LH as an underlying tonal melody. Of particular interest is the other type: rising tones may also result from *tone trapping* (Tranel 1992/1994: 112). That is, contour tones are created when a tone that is set free due to glide formation re-associates to another TBU, or the Stranded Tone Principle discussed in §1.2.1. Of central interest is the observation that the Stranded Tone Principle does not apply across the board in Margi: Once glide formation occurs, rising tones are created *within a word* but are banned *across words* (i.e. only high tone survives, to which I will return shortly). Some representative examples will be presented in the following data section.

3.3.2 Conditions on Rising Tone Formation

In Margi, some instances of tonal modifications arising from vowel fusion or glide formation are nicely captured by Clements and Ford's (1979) Stranded Tone Principle, as exemplified by the following definite-suffixed forms. In (37)a, the low tone set afloat by gliding re-associates to the initial vowel of the definite suffix, resulting in a rising tone, while the high tone in the definite suffix links to the preceding stem vowel in (37)b. It is evident that the tone that is set afloat re-associates to the vowel that conditions fusion.[7]

---

[7] Regarding low-toned disyllabic stems, rising tones are not attested in definite suffixation. A common response to this puzzle is that the unsuffixed words in (i) have one underlying L tone, multiply linked to both syllables (Tranel 1992/1994, Kenstowicz 1994). This will not concern us here because the fact shows that rising tones are simplified

(37) Glide formation/Vowel deletion in árì-suffixation

a. hù-árì	→	hwǎrì	‘grave’
L-H.L		L͡H.L	
b. cédè-árì	→	céděrì	‘money’
H.L-H.L		H.L͡H.L	
c. àdíkò-árì	→	àdíkwǎrì	‘kerchief’
L.H.L-H.L		L.H.L͡H.L	

Let us now turn to the context in which the purported rising tones fail to surface. The phenomenon is illustrated with the *á*-genitive construction. In this construction, the possessed noun precedes the genitive affix *á* (GEN). When appropriate, vowel fusion and glide formation also take place as hiatus resolution. As we can see, no rising tone is created. Vowel fusion occurs in (38)a: the genitive particle *á* merges with the preceding vowel /a/ and the high tone of this particle is retained on the word *ŋwa* ‘face’ whose underlying tone is L, suggesting that a rising tone was once created at a stage of derivation and was reduced to a high level tone on the surface. Likewise, we see in (38)b that glide formation takes place and again no rising tone occurs. Following Tranel (1992/1994), I assume the á-genitive construction has the following structure: [$_{DP}$ *tsì* (hand) [$_{PP}$[$_{P}$ *á* (GEN) [$_{DP}$ *wù* (tree)]]]]. This is primarily because Margi is a head-initial language. Therefore, it is not unproblematic to treat the particle *á* as a suffix to the possessed noun (contra Hoffmann 1963).

---

when preceding a L tone: L.LH → L.H. Tembo (Bantu) also has the same phonotactic restriction (Kaji 1996).

(i) a. làgù	→	làgwárì	‘road’
b. màlà	→	màlárì	‘woman’

(38) The á-genitive construction ([possessed] [á+possessor])

a. [ŋwà][á+bzə́r] → ŋwá bzə́r '[face] [GEN+boy]: the boy's face'
L H H → H H

b. [tsì][á+wù] → tsyá wù '[hand] [GEN+tree]: branch'
L H L → H L

Is the absence of rising tone across word boundary attributable to post-lexical phonology? Post-lexical rules are typically about allophonic changes but rising tone formation *is* attested in suffixation. The phonetic motivations for post-lexical processes are normally transparent but Margi rising tone is licensed even in non-final position (cf. (37)). So duration is not at issue. On the other hand, monosyllabic words can bear rising tone, indicating that word-final position is compatible with rising tone. Finally, penultimate lengthening should be irrelevant, too. First, the examples under examination are monosyllabic words. Second, this phenomenon is not documented in the relevant literature, as far as I know of.

All in all, it seems that the present asymmetry boils down to the issue of word boundary: rising tone is not created across words. It is important to note that this generalization is in accordance with the lower part of the proposed ranking: RELCORR$_{WD}$ » RELCORR. Word-internal contour preservation is more rigid than contour preservation across word boundary. In the word domain, rising tone is licensed by the word-internal contour correspondence constraints but this is not the case in the larger domain.

Before we launch into the analysis, let us first discuss what kind of problem the standard OT analysis encounters for the case at hand, to which I turn in the following section.

3.3.3 A Ranking Paradox

In this section, I consider the standard OT account for the asymmetry in question. First, I assume that the glide formation or vowel fusion is motivated by the top-ranked constraint *VV 'Immediately adjacent vowels are disallowed.' In order to ensure that the tone that is set free due to glide formation re-links to a TBU, we have to rank Max-(H) and Max-(L) over *Rise. As we expect, candidate (a) is selected as the optimal output in the following tableau.

(39) Rising tone is licensed within a word: hù-árì → hwǎrì 'grave-definite suffix'(37)b

hù-árì L-HL	*VV	Max-(H)	Max-(L)	*Rise
☞ a. hwǎrì LH.L				*
b. hwárì H.L			*!	
c. hwàrì L.L		*!		
d. hùárì L.H.L	*!			

On the other hand, since contour tones are banned across word boundary, we have to rank *Rise over Max-(L). But it is obvious that this new ranking below fails to predict the correct output in tableau (40). Conversely, given the ranking in (40), it is predicted that rising tone should appear without any problem across words.

(40) No rising tone across words: (38)b

[tsì] [á+wù]	MAX-(H)	*RISE	MAX-(L)
☞ a. [tsyá] [wù] H L			*
b. [tsyǎ] [wù] LH L		*!	

([tsì] [á+wù] → [tsyá] [wù] '[hand] [GEN+tree]: branch')

Furthermore, the new ranking in (40) also wrongly predicts that an underlying rising tone would never surface.

(41) Lexically specified rising tone: *və̌l* 'fly; jump'

və̌l LH	MAX-(H)	*RISE	MAX-(L)
☹ a. və̌l		*!	
b. və̀l	*!		
!☞ c. və́l			*

We should also consider the possibility that underlyingly present rising tone is able to surface but a "newly created" rising tone is ruled out. This opaque derivation can be analyzed in McCarthy's (2002) Comparative Markedness. Since rising tone formation creates the new marked structure, i.e. candidates which have been altered by GEN, $_{N}$NO-RISE may be active in eliminating rising tone formation. The main problem with this approach is that $_{N}$NO-RISE fails to distinguish 'word-internal' rising tone formation from rising tone resulting from two tones across words. These two types of newly created rising tone should be both recognized as new marked structures.

Finally, invoking Output-Output faithfulness (Benua 1997) seems hopeless, too. If we postulate that the genitive particle á is a suffix to the preceding noun, then the Output-Output relation can be established. Recall [tsì] [á+wù] → [tsyá] [wù] '[hand] [GEN+tree]: branch' The surface tone of the fused form *tsyá* is from the underlying high tone of the genitive affix á, suggesting that the IDENT-OO-(TONE) is also violated between the 'base' form *tsì* (low tone) and the 'derived' form *tsyá* (high tone). In addition, the hypothetical form *tsyǎ* may be more faithful than the actual form *tsyá* because *tsì* and *tsyǎ* both have a low tone.

In sum, I have shown that a markedness account is problematic in that rising tone should be allowed or disallowed across the board. Furthermore, although comparative markedness is useful in eliminating new marked structure, this approach also prevents word-internal rising tone formation. Finally, if we resort to OO-faithfulness, the actual output form also incurs violation of IDENT-OO-(TONE). The Margi data pose a problematic case for the standard OT approach. In view of contour correspondence, as we have mentioned earlier, this asymmetry is again an instantiation of 'degree of proximity' discussed at the outset. The analysis is provided in the following section.

3.3.4 Why Word Boundary Matters

Our analysis begins with the simplest case: How does lexically specified rising tone surface in the first place? This is straightforwardly explained by ranking RELCORR$(x<y)_{\text{NUC/WORD}}$ over *RISE. Notice also that the nucleus- and word-internal contour correspondence constraints are indistinguishable here because the input is a monosyllabic word.

(42) Underlying LH survives on the surface

və̌l LH	RELCORR $(x<y)_{\text{NUC}}$	RELCORR $(x<y)_{\text{WD}}$	*RISE
☞ a. və̌l			*
b. və̀l	*!	*	
c. və́l	*!	*	

Violation profile of the candidates w.r.t. RELCORR $(x<y)_{\text{NUC/WD}}$

Input	Output	Mapping	Violation
$l_x h_y$	$l_a h_b$	x<y → a<b	No
$l_x h_y$	$l_a l_b$	x<y → a=b	Yes
$l_x h_y$	$h_a h_b$	x<y → a=b	Yes

From the above tableau, it should be clear that only the winning candidate (a) preserves the underlying LH contour: the 'smaller than' relation ($x<y$) is faithfully realized in the output. In the same vein, rising tones appear without any problem provided that word-internal contour preservation constraint outranks *RISE. In the following tableau, only the parenthesized parts (i.e. the site of glide formation) are under evaluation. Notice also that RELCORR$(x=y)_{\text{NUC}}$ must be inactive. Otherwise, rising tone formation will not be possible.

(43) Rising tone resulting from glide formation

(hù-á)rì L-HL	RELCORR $(x<y)_{\text{NUC}}$	RELCORR $(x<y)_{\text{WD}}$	*RISE	RELCORR $(x=y)_{\text{NUC}}$
☞ a. (hwǎ)rì (LH).L			*	*
b. (hwá)rì (H).L		*!		
c. (hwà)rì (L).L		*!		

(hù-árì → hwǎrì 'grave-definite suffix')

Violation profile of the candidates (the parenthesized) w.r.t. RELCORR $(x<y)_{WD}$

Input	Output	Mapping	Violation
$(l_x h_y)$	$(l_x h_y)$	$x<y \rightarrow a<b$	No
$(l_x h_y)$	$(h_x h_y)$	$x<y \rightarrow a=b$	Yes
$(l_x h_y)$	$(l_x l_y)$	$x<y \rightarrow a=b$	Yes

Violation profile of the candidates (=vowel /a/ in árì) w.r.t. RELCORR $(x=y)_{NUC}$

Input	Output	Mapping	Violation
$(h_x h_y)$	$(l_x h_y)$	$x=y \rightarrow a<b$	Yes
$(h_x h_y)$	$(h_x h_y)$	$x=y \rightarrow a=b$	No
$(h_x h_y)$	$(l_x l_y)$	$x=y \rightarrow a=b$	No

We see in candidate (a) that the rising tone resulting from glide formation survives because RELCORR$(x>y)_{WD}$ dominates *RISE. The L on the stem *hu* and the H on the initial syllable *a* of the definite suffix *ari* stand in relational correspondence: $x>y$ within the word domain. Here the markedness constraint *RISE is overridden by the requirement of contour preservation. By contrast, candidates (b-c) lose out because the active word-internal relational correspondence constraint RELCORR$(x>y)_{WD}$ rules out candidates that undergo contour simplification. I.e. the 'greater than' relation is not supposed to be tampered in the output. In summary, the present analysis predicts that rising tone, be it lexical or derived, should be faithfully rendered within a word. As we have seen, this prediction is borne out.

With regard to the 'ill-formedness' of rising tone across word boundary, the analysis proceeds as follows. If the most general contour preservation constraint RELCORR$(x<y)$ is dominated by *RISE, no contour tone is formed. Moreover, MAX-(H) must outrank MAX-(L) so that the "free" high tone survives on the surface.

(44) [tsì] [á+wù] → [tsyâ] [wù] 'hand+GEN+tree: branch'

[tsì] [á+wù] [L] [H L]	RELCORR $(x<y)_{\text{NUC/WD}}$	MAX-(H)	*RISE	MAX-(L)	RELCORR $(x<y)$
☞ a. [tsyá] [wù] [H] [L]				*	*
b. [tsyǎ] [wù] [LH] [L]			*!		
c. [tsyà] [wù] [L] [L]		*!			

(Where []=word boundary)

In Margi, it is evident that the Stranded Tone Principle applies on the basis of degree of proximity between the free tone and its host. This conditioning factor is characterized by the partial ranking: RELCORR$_{\text{WD}}$ » RELCORR.

3.4 Discussion and Conclusion

In this chapter, I have proposed that contour correspondence constraints should be relativized to different contexts. This contextualization is based on the speaker's presumed knowledge according to which a contour change that is perceptually more salient is banned before a contour change that is less so. This assumption was translated into the intrinsic ranking: RELCORR$_{\text{NUC}}$ » RELCORR$_{\text{WD}}$ » RELCORR. I have also demonstrated that tone sandhi in San Miguel el Grande Mixtec and Margi provide evidence in support of the upper and lower part of this ranking, respectively.

Returning to the comparison with Dilley's (2004) relational features, it is fair to say that using relational features makes wrong predictions in the data we have discussed in the preceding sections. For example, an HL sequence within a syllable and across a syllable boundary is indistinguishable in Dilley's tone interval. But we have seen that contour preservation is sensitive to the degree of different cohesiveness. It is conceivable that IDENT-(relational features) may also be contextualized in the same way we have employed in this chapter. A more serious problem, however, lies in the fact that relational features should also be regulated by, for example, the anti-insertion constraint DEP. If DEP-(relational feaures) is inactive, then one of the undesirable results is that an LH melody is inserted in a polysyllabic word. This possibility is never attested, to the best of my knowledge. In contrast, preservation of the phonological height relation is not featural specification. Relational correspondence is rooted in similarity in phonology.

Chapter 4 On the (Non-)Contouricity Agreement

In chapters 2 and 3, I have discussed and demonstrated three of the four relations proposed in chapter 2, namely, "greater than" ($x>y$), "smaller than" ($x<y$) and "equal to" ($x=y$), and interactions thereof. The goal of this chapter is to validate the fourth relation, i.e. the "non-equal to" relation ($x\neq y$). In terms of similarity in phonology (cf. Steriade's (2001) P-map approach), the non-equal to relation is at first blush considerably *counterintuitive* because under this relation, mappings such as $\overset{\frown}{\mathrm{HL}} \rightarrow \overset{\frown}{\mathrm{LH}}$ are treated as being more faithful (or more precisely, perceptually less deviant) than mappings like $\overset{\frown}{\mathrm{HL}} \rightarrow \mathrm{H}$. The reason is as follows: Supposing that Language *L* contrasts three tone levels, H, M and L, for an underlying steep falling tone $\overset{\frown}{\mathrm{HL}}$, the most faithful correspondent in the output must be a steep falling tone $\overset{\frown}{\mathrm{HL}}$, too, because no discernable deviation ever occurs in this mapping. The second-best match should be a smooth falling tone $\overset{\frown}{\mathrm{ML}}$.[1] This is because both $\overset{\frown}{\mathrm{HL}}$ and $\overset{\frown}{\mathrm{ML}}$ have a fall in pitch. In terms of contour correspondence, we can say that the two falling tones both manifest the greater than relation ($x>y$), hence are similar. All else being equal, the only difference lies in steepness, $\overset{\frown}{\mathrm{HL}}$ being steeper and $\overset{\frown}{\mathrm{ML}}$ shallower in slope.[2] So it should not be problematic to say that $\overset{\frown}{\mathrm{ML}}$ deviates the least from the input $\overset{\frown}{\mathrm{HL}}$, provided that the output is not $\overset{\frown}{\mathrm{HL}}$ itself. As for output correspondents without a falling

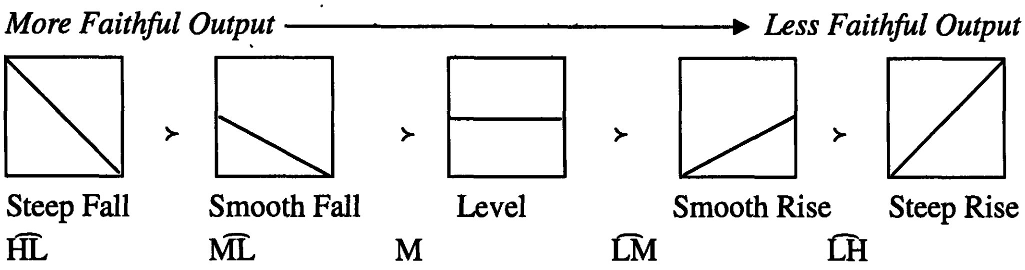
---

[1] Another possibility is $\overset{\frown}{\mathrm{HM}}$. Let us simply ignore it here.

[2] See chapter 5 for more discussion on cases in which slope difference is crucial.

contour, it is conceivable that level tones, rather than rising tones, should be more similar to falling tones. Gandour (1978), citing the experimental results from Gandour and Harshman (1978), claims that "direction" is one of the five dimensions relevant for the perception of contour tones. Thus, it should be safe to say that mappings of the opposite contour direction are least favorable in terms of perceptual similarity. The present discussion is graphically presented as follows. Notice that the same rationale also applies to cases in which the input is any other type's contour tone.

(1) Degree of faithfulness of output correspondents for a steep falling tone ($\widehat{\mathrm{HL}}$)



(Where ≻ means "one step away" in the present hypothetical perceptual scale)

From (1), mappings such as $\widehat{\mathrm{HL}} \rightarrow \widehat{\mathrm{LH}}$ or $\widehat{\mathrm{LH}} \rightarrow \widehat{\mathrm{HL}}$ are doomed if similarity plays a decisive role here. Given a contour tone in the input, it is clear that level tones are, by transitivity, a more faithful output correspondent than a contour tone with the reversed contour direction. In the tone sandhi literature, however, Input-Output mappings of two dissimilar contour tones (i.e. of the reversed contour direction) are not uncommon. Tonal alternations of this sort are often labeled as "contour metathesis" (Bao 1990, 1999), or "contour dissimilation" (Chen 2000) in the Chinese tone sandhi literature and are conventionally analyzed in the following fashion:

(2)
$$\begin{array}{cccc} \text{CTU} & & \text{CTU} & \\ \wedge & \rightarrow & \wedge & \\ \text{HL} & & \text{LH} & \\ \text{LH} & & \text{HL} & \text{(where CTU = contour tone unit)} \end{array}$$

It seems that employing the "non-equal to" relation may not be empirically distinguishable from the CTU-based approach. The goal of this chapter is to tease apart the "non-equal to" relation and other alternatives. The empirical evidence comes from a phenomenon I termed "*contouricity agreement*" in Shaoxing Chinese disyllabic tone sandhi (Ping 2001a). By "contouricity," I mean the *contour/level* distinction. (Non-)contouricity agreement thus refers to the phenomenon whereby adjacent tones must be either contour tones or level tones in a disyllabic domain. For ease of exposition, I would like to present the crux of the argument with some representative Shaoxing Chinese tone sandhi data in the following section.

4.1 First Approximation: Some Representative Data

Our discussion begins with an example of (non-)contouricity agreement in the abstract. Suppose that Language *L* has the following phonotactic restrictions on tone concatenation in a disyllabic domain (3). If the first syllable carries a contour tone, then the second syllable must be also specified with a contour tone (3)a-b. Conversely, in (3)c-d, if the initial tone is a level tone, then the final syllable must be a level tone, too.

(3) Hypothetical phonotactic restrictions in Language *L*

a. HL-HL, HL-LH, LH-LH, LH-HL
b. *HL-H, *HL-L, *LH-H, *LH-L
c. H-H, L-L, H-L, L-H
d. *H-HL, *H-LH, *L-HL, *L-LH

In terms of contour correspondence, the phenomenon in question can be understood as follows. For (4)a, we can say that the non-equal to relation is motivated by an affinity shared among contour tones, i.e. rising and falling tones may be similar *because their initial and the final tones are not identical*, .e.g. L≠H and H≠L, or more precisely, the tone values associated with the temporal spans. The very affinity has been labeled as contouricity in the preceding section. With regard to (4)b, the equal to relation can be likewise motivated.

(4) The contour/level distinction in terms of relational correspondence

e. Contour tones exhibit the "non-equal to" relation ($x \neq y$).
f. Level tones exhibit the "equal to" relation ($x = y$).

From the preceding discussion, the phonotactic restrictions in (3) can be analyzed in the following way. First, let us assume that the first tone and the second tone must *agree* in (non-)contouricity. This requirement can be formalized via the AGREE(F) approach (Lombardi 1996, 1999, Bakovic 2000),[3] if we subscribe to the view that the contour/level distinction may be grounded on the affinities shared among contour/level tones. Or more precisely, the phonotactic restrictions in (3) are interpreted as a requirement to enhance the similarities regarding (non-)contouricity between the first and the second tone. Armed with these assumptions, I formally define two AGREE constraints in (5) and (6).

---

[3] Rose and Walker's (2004) "Agreement by Correspondence" approach may not be suitable for tone sandhi (see also (Hansson 2001)). Tone spreads only via autosegmental linking, to the best of my knowledge.

(5) AGREE-T_1T_2 ($x \neq y$)

Let T_1 be a tone contained within Rime R_i.
Let t_1 be a tone value contained within Rime R_i. Let S_1 be a temporal span associated with t_1
Let t_2 be a tone value contained within Rime R_i. Let S_2 be a temporal span associated with t_2.
S_1 precedes S_2.

Let T_2 a tone contained within Rime R_j.
Let t_3 be a tone value contained within Rime R_j. Let S_3 be a temporal span associated with t_3
Let t_4 be a tone value contained within Rime R_j. Let S_4 be a temporal span associated with t_4.
S_3 precedes S_4.

T_1 and T_2 are adjacent tones. T_1 precedes T_2.

Let $t_1 = x$, $t_2 = y$, $t_3 = a$, and $t_4 = b$.

If $x \neq y$, then $a \neq b$.

(6) AGREE-T_1T_2 ($x=y$)

Let T_1 be a tone contained within Rime R_i.
Let t_1 be a tone value contained within Rime R_i. Let S_1 be a temporal span associated with t_1
Let t_2 be a tone value contained within Rime R_i. Let S_2 be a temporal span associated with t_2.
S_1 precedes S_2.

Let T_2 a tone contained within Rime R_j.
Let t_3 be a tone value contained within Rime R_j. Let S_3 be a temporal span associated with t_3
Let t_4 be a tone value contained within Rime R_j. Let S_4 be a temporal span associated with t_4.
S_3 precedes S_4.

T_1 and T_2 are adjacent tones. T_1 precedes T_2.

Let $t_1 = x$, $t_2 = y$, $t_3 = a$, and $t_4 = b$.

If $x = y$, then $a = b$.

The above constraints motivate tone extension in an effort to avoid tone sequences that are *not* similar in terms of (non-)contouricity, provided that unit correspondence constraints like IDENT-(Tone) are dominated in (7).

(7) AGREE-$T_1T_2(x\neq y)$ // AGREE-$T_1T_2(x=y)$ » IDENT-(Tone), MAX-(Tone), DEP-(Tone), etc.

Given the above ranking, it is predicted that a contour tone will always map to a contour tone in the output. As shown in (8), given that the falling contour of the initial tone 31 is faithfully rendered in the output, its following tones must be either a falling or a rising tone. Under the pressure from AGREE-$T_1T_2(x\neq y)$, level tones must map to a contour tone. This is evidenced in (8)c, where the following tone 55 changes to contour tone 34 when preceded by contour tone 31.

(8) Some faithful/unfaithful sequences by virtue of AGREE-$T_1T_2(x\neq y)$ in Shaoxing Chinese

	Input	Actual output: $(x\neq y)$-$(a\neq b)$	Possible output *$(x\neq y)$-$(x=y)$
a.	31-223	32-23	*32-22
b.	31-31	32-21	*32-22
c.	31-55	32-34	*32-22, *32-33 , etc.

(Note that level tones are only those with the identical juxtaposed digits. See §4.4 for more details. Checked tones are underlined throughout.)

Likewise, the active AGREE-$T_1T_2(x=y)$ constraint requires that a level tone must be followed by a level tone. Some representative examples in Shaoxing Chinese are given below. As we can see, contour tones are flattened when following level tone 55. Notice also that the citation checked tone 55 is lowered to 33 in initial sandhi position.

(9) Some faithful/unfaithful sequences by virtue of AGREE-$T_1T_2(x=y)$ in Shaoxing Chinese

	UR	Actual SR: $(x=y)$-$(a=b)$	Possible SR: $(x=y)$-$(a\neq b)$
a.	55-31	33-33	*33-31, *33-21, etc.
b.	55-223	33-33	*33-23

(Note that level tones are only those with the identical juxtaposed digits. See §4.4 for more details. Checked tones are underlined throughout.)

It is important to note that the above data pose a severe challenge to unit correspondence and markedness accounts. In (10)a-b, 31 has two sandhi forms in final position, 21 and 33. In particular, the mapping "**31 → 33**" violates every relevant unit correspondence constraints, MAX (i.e. the boldfaced 1 is deleted), DEP (i.e. the boldfaced 3 is inserted), and finally, IDENT (i.e. featural identity is inexact). On the other hand, markedness is difficult to account for the alternation in which a less "marked" level checked tone 55 becomes a more "marked" checked rising tone 34 in (10)c. So the data cast doubt on a duration-based account for contour tone licensing (Gordon 2001, Zhang 2002b, inter alia).

(10) a. 31-31 → 32-21
b. 55-**31** → 33-**33**
c. 31-55 → 32-34

In contrast, I have demonstrated that the desired results are more straightforwardly obtained by appeal to the (non-)equal to relations with the AGREE approach. The rest of this chapter is devoted to a comprehensive illustration and detailed examination of the phonetics and phonology of Shaoxing Chinese tone sandhi. I will argue that a proper treatment of the phenomena in question cannot be attained unless the (non-)equal to relations are factored in the analysis.

4.2 The Citation Tone Inventory

Shaoxing Chinese (henceforth SX), a dialect of Northern Wu Chinese, is spoken in Zhenjiang province, China (Ping 2001a, Zhang 2006). This section is primarily concerned with the citation tone inventory of SX. Some phonetic details of interest, including the phonation-tone correlation, the F_0 curves and rime duration for each citation tone, are presented.

4.2.1 Tone and Phonation Register

SX has eight tones in isolation: six long (or non-checked) tones (i.e. tones on sonorant-final syllables) and two checked tones (i.e. tones on glottal stop-terminating syllables). Each of the four Middle Chinese tones (i.e. *ping* 'even', *shang* 'ascending', *qu* 'departing' and *ru* 'entering') is split neatly into a H- and a L-register, yielding a symmetrical eight-tone system. The pitch value and averaged duration for each of the eight citation tones are illustrated below. Unless otherwise noted, all data are taken from Ping's (2001a) acoustic study.[4]

(11) The citation tone inventory of Shaoxing Chinese

a. Long tones (sonorant-final syllable)

H-register	Duration	SD	N	L-register	Duration	SD	N
Tone 1 551	168.5	22.4	15	Tone 2 231	239.3	22.2	15
Tone 3 334	384.5	33.0	15	Tone 4 223	430.0	44.5	15
Tone 5 31	366.5	40.7	15	Tone 6 221	383.7	40.7	15

b. Checked tones (glottal stop-terminating syllable)

H-register	Rime	SD	N	L-register	Rime	SD	N
Tone 7 $\underline{55}$	89.2	25.6	15	Tone 8 $\underline{23}$	127.1	42.2	15

(Pitch values are transcribed at the five-point scale (i.e. Chao's letters), where 5 is the highest and 1 is the lowest; Rime duration is given in ms; SD=Standard Deviation; N=number of tokens; checked tones are underlined throughout.)

From (11), we can see that each H-register tone has its L-register counterpart and the paired tones basically match in their contour shape, except the checked tones (Tone 7/8). For example, we can say that Tone 3 (334) is an H-register rising tone and its L-register counterpart, Tone 4 (223) is an L-register rising tone. Of interest is the fact that L-register tones are slightly longer

---

[4] The recording was made by one male native speaker HT. The wordlist was given in Chinese characters and was read in isolation (Ping 2001: 101-102). As far as I can tell, those recorded monosyllables are basically open syllables with the low vowel /a/. The acoustic measurements were conducted with KAY CSL 4300B. See also Zhang (2006) for a recent OT analysis of SX tone sandhi.

than H-register tones in a consistent manner (albeit statistically not significant: t(2)=-2.88, $p=0.1$). This fact accords with Maddieson's (1978) proposal regarding universals of tone according to which low-toned vowels tend to be longer than high-toned vowels. The lengthened duration of the L-register syllables can also be interpreted as a by-product of what Silverman (2002) termed "laryngeally complex" vowel. That is, an L-register vowel in SX can be viewed as a concatenation of a breathy vowel followed by a modal vowel: [V̤V].[5] It has been instrumentally confirmed that such "register" contrasts are tightly related to phonation differences in most varieties of Wu Chinese.[6] More specifically, L-register tones occur only on syllables with breathy voice (or murmur; I shall use them interchangeably), whereas H-register tones cooccur with 'plain' syllables, i.e. syllables with modal (or clear) phonation. This contrast is rooted in a well-known fact: breathy voice lowers F_0 (Gordon and Ladefoged 2001, Silverman 2002, Hombert et al. 1979, Hombert 1978). Turning back to Wu Chinese, Cao and Maddieson (1992), Ren (1992) and Zhu (1999) report that breathy voice is most salient in the vocalic onset and fades away before the middle point of a vowel in a handful of Wu Chinese dialects they investigated (see fn. 6). I assume that SX has these properties as well because the phonation-tone interaction is robustly attested in this language as well. Consider now (12). Notably, the phonation contrast determines the surface pitch range of the tonal onset. Supposing that 3 is the lower limit of a low tone in modal phonation, we can see that breathy voice-affected tones consistently start at 2.

---

[5] Silverman (2002) further argues that pitch is not well cued on the breathy portion so that the modal portion is lengthened to accommodate tonal contours.

[6] These Wu Chinese dialects include: *Changyinsha Chinese*: Cao and Maddieson 1992; *Ningbo Chinese*: Cao and Maddieson 1992; *Shanghai Chinese*: Cao and Maddieson 1992, Ren 1992, Zhu 1999; *Suzhou Chinese*: Shi 1983, (Iwata et al. 1991); *Zhenhai Chinese*: Rose 1981; *Wenzhou Chinese*: Cao and Maddieson 1992.

(12) The Phonation-Tone Correlation

Phonation	Tone	Pitch value	Syllable type
Modal	Tone 1	551	CVV/CVN (C = p, p^h, m, ∅, etc.)
	Tone 3	334	
	Tone 5	31	
	Tone 7	55	
Breathy	Tone 2	231	C̤V̤V/C̤V̤N (C = p, m, ∅, etc.)
	Tone 4	223	
	Tone 6	21	
	Tone 8	23	

As a caveat, before the section is closed, we have to distinguish "phonation register" from "tonal register" (Yip 1993). Tonal register, or H- and L-register, refers to tones settled in the higher/lower part of the F_0 range. By contrast, phonation register in SX contrasts "modal" and "breathy" voice. Furthermore, it has also been noted that the phonation distinction in Wu Chinese is restricted to word-initial position (Cao and Maddieson 1992, Ren 1992, Zhu 1999, among many others).

(13) Tonal register vs. Phonation

	Pitch (=Tonal register)	Position
Modal	Higher and Lower Pitch	All
Breathy	Lower pitch	Word-initial

In sum, it is important to bear in mind that breathy voice only occurs with L-register tones but modal phonation is compatible with both H- and L-register tones. Throughout the discussion, I will disambiguate the use of "register" if necessary.

Let us now turn to look at the normalized F_0 curves and phonetic lengths for each tone in the following sections.

4.2.2 Acoustic Properties of Long Tones

In this section, some essential phonetic traits of SX long tones are described, in particular, F_0 and duration. All of the data are replotted from Ping (2001). The normalized F_0 trajectories are plotted against normalized time. The normalized points for each long tone are those at 0%, 5%, 20%, 40%, 60%, 80% and 100%. As a notational convention throughout this chapter, H-register tone is represented with filled shape and L-register tone hollow shape throughout this chapter.

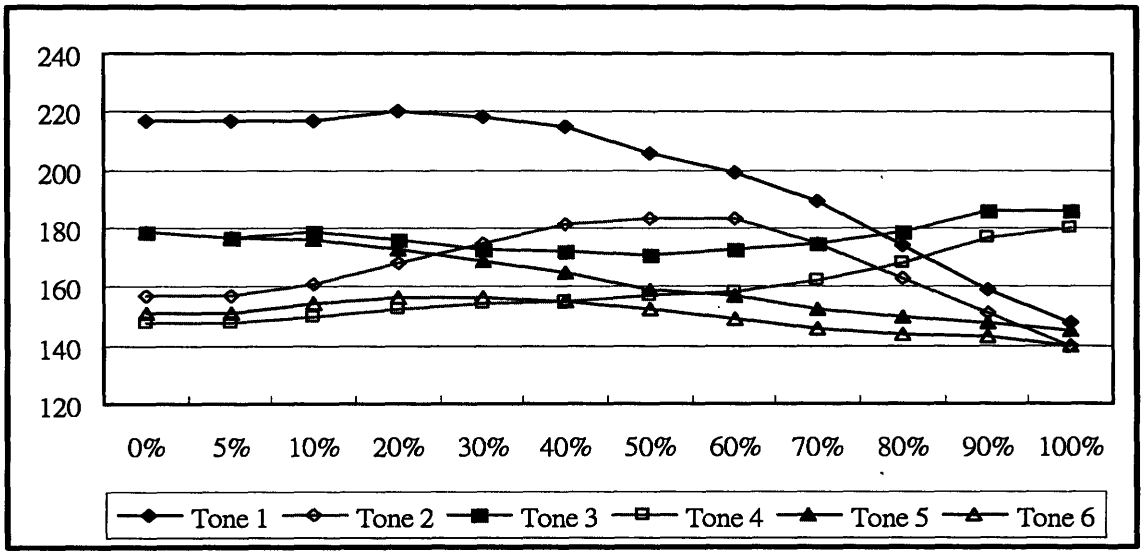


Figure 4-1 Normalized F_0 for long tones in Shaoxing Chinese

The normalized F_0 values were in turn converted into logarithmic Z-score (henceforth LZ(-score), Zhu 1999). The timepoints of the LZ normalized F_0 values are given at 5%, 20%, 40%, 60%, 80% and 100%.[7] LZ normalized values are the primary method to compare the F_0 contours among tones (see also §1.3.1).

---

[7] The F_0 values at the 0% point for all tones and at the 100% point for Tone 1 (551) and Tone 2 (232) were not provided because of relatively large standard deviations (Ping 2001).

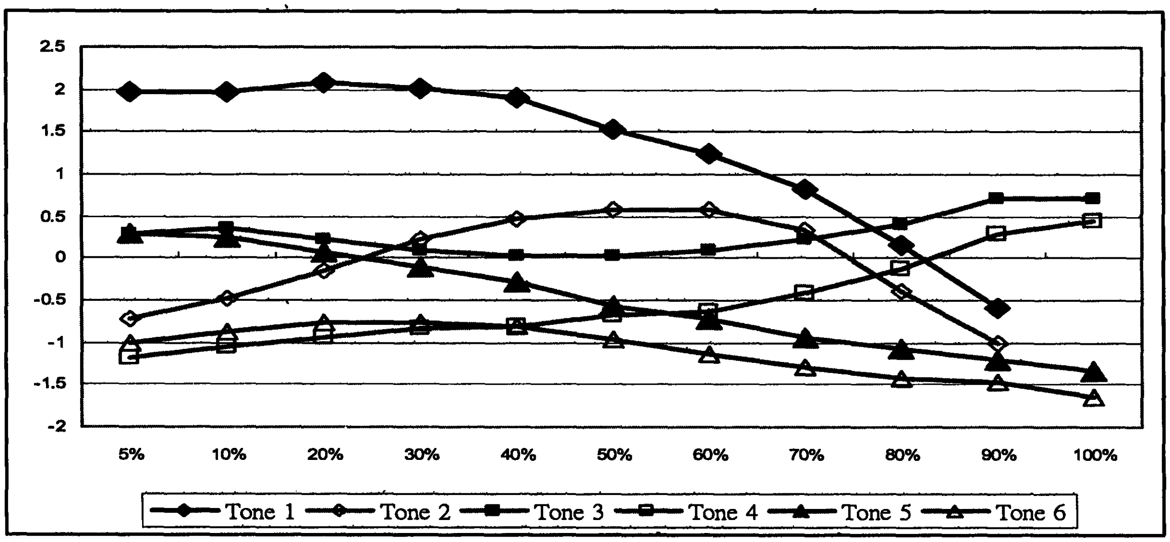


Figure 4-2 LZ Normalized F_0 for long tones in Shaoxing Chinese

There are several things to note about the above figures. First, the register distinction is very salient in the beginning. All the starting points of H-register tones are above 0 in LZ or around 180 Hz, whereas those of L-register tones consistently start at about -1 in LZ or 150 Hz, suggesting that the F_0. lowering effect of murmur is quite robust. The register distinction is neutralized at the ending points. Since the ending points (and the second half) of the paired tones are very close to one other, and since the initial F_0 contour is predictable by the presence or absence of murmur, these facts suggest an *allotone* analysis of the paired tones (i.e. Tone 1/2 are allotones and so are Tone 3/4 and Tone 5/6).

Second, Tone 1 (551) can be described as a high falling tone. This tone starts at 2 (in LZ) and the contour stays level till the 40% point. Then we see a steep fall. Its L-register counterpart, Tone 2 (231), looks like a convex tone at first sight. As mentioned earlier, this initial F_0 lowering is caused by murmur. The pitch contour reaches its initial target at the 30% point or a bit earlier (i.e. somewhere around 0 in LZ). Tone 2 (231) thus can be regarded as a mid-falling tone with the lowered F_0 onset due to murmur.

Third, it appears that the overall contour of Tone 3 (334) is substantially flat. This tone has a tiny dip in the first 30% of normalized duration. The dipping onset might be attributable to the following production effect: In absence of time pressure, speakers may firstly start with the neutral pitch level (i.e. the midrange: somewhere around 0.5 in LZ) and then the contour slides down to reach the initial low target. As we will see, initial declines of this sort disappear in sandhi environments, presumably due to greater time pressure. So it should be safe to say that this dip is not linguistically relevant. As for the L-register rising tone, Tone 4 (223) has a greater pitch excursion: the difference between the starting and the ending normalization points is approximately 2.3 in LZ, markedly greater than that of Tone 3 (334), 0.4.

Fourth, Tone 5 (31) and Tone 6 (221) can be described as mid-falling (without the lowered F_0 onset) and low-falling tones, respectively. Again, we can see that the first 30% of the F_0 contour is slightly compressed for Tone 6 (221). Again, this is attributable to the effect of murmur. Thus far, we have discussed the F_0 curves for each of the long tones in some detail. The absolute duration for each long tone are plotted below. The error bar indicates one standard deviation.

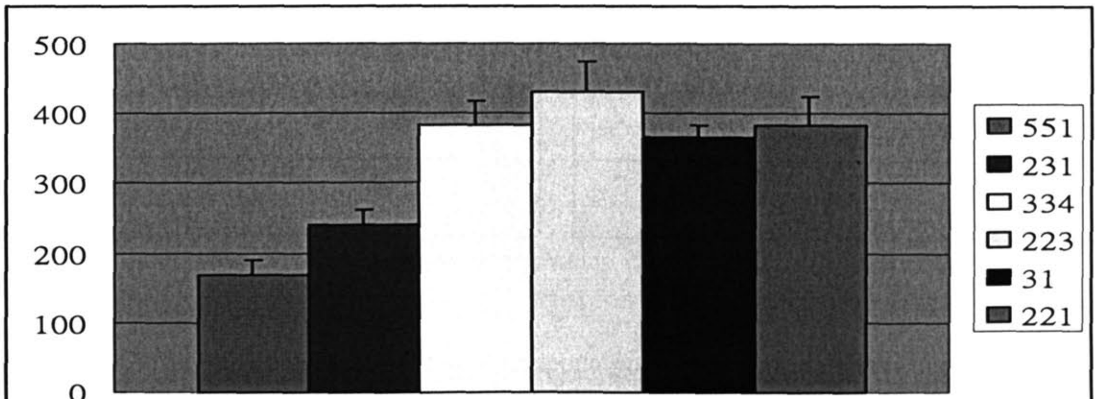


Figure 4-3 Duration (in ms) for long tones in Shaoxing Chinese

This section is closed by raising a puzzling observation: as shown in Figure 4-3, Tone 1 (551) and

Tone 2 (231) are significantly shorter than the other long non-checked tones, suggesting that Tone 1/2 have some peculiarity. As we will see, Tone 1/2 do have anomalous behaviors in disyllabic tone sandhi.

4.2.3 Acoustic Properties of Checked Tones

SX has two checked tones. The corresponding (LZ-)normalized F_0 curves are plotted below. Tone 7 (55) is a high level checked tone. This is the only level tone among citation tones, according to the criteria that will be provided in the following section. As for the L-register checked tone (Tone 8), it is a rising tone, transcribed as 23. As we can see, the starting pitch levels of the two checked tones are quite distinct, again due to different phonation registers. Unlike the long tones, however, the final points of the two checked tones do not end at a similar point. They are separated by about 1 in LZ.

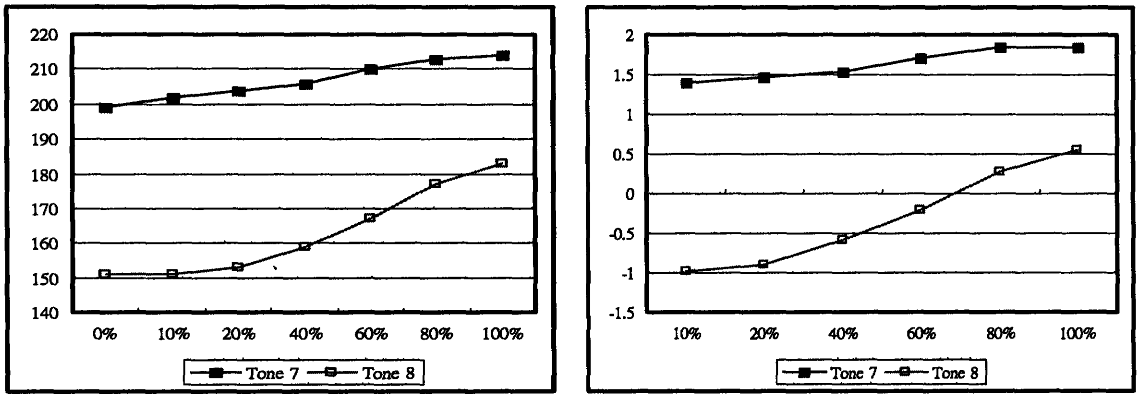


Figure 4-4 Normalized F_0 (left) and LZ normalized F_0 (right) for checked tones

As for duration, it is evident that the rising checked tone (Tone 8) is considerably longer than the high level checked tone (Tone 7) by around 40 ms. It has been well-established that it takes more time to implement a rising tone (Ohala and Ewan 1973, Sundberg 1973, 1979, Ohala 1978, Xu and Sun 2002). The present data are thus regarded as yet another instantiation of this well-known fact.

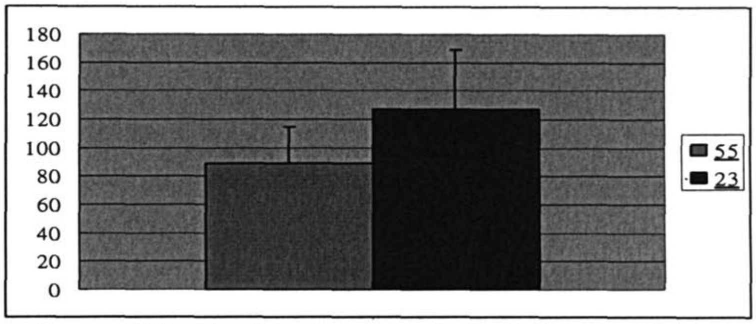


Figure 4-5 Duration (in ms) of checked tones in Shaoxing Chinese

This completes the phonetic description of SX citation tones. Some discussion is accordingly given in the next section.

4.2.4 Discussion

Having discussed the citation tones in terms of (LZ) normalized F_0 and duration, I raise three points of interest in this section. The first point concerns the contour/level distinction. As we have seen, the overall contour of Tone 3 is substantially flat. There is nevertheless convincing evidence in support of the claim that Tone 3 is underlyingly a rising tone. For the present purpose, we can take the F_0 fluctuation on the edges as the threshold value for a contour tone. Recall that Tone 3 has an initial dip due to breathy voice so the 20% point (=0.09) is chosen as the starting point and the 100% point (=0.71) the endpoint. Judging from the fact that the difference of the initial (20%) and final (100%) points in LZ-score (henceforth Δ_{LZ}) is greater than 0.5, I suggest that contour tones in SX should be tentatively defined as follows. LZ scores are accordingly converted into the five-point scale in the table (15).

(14) Definition of a contour tone in Shaoxing Chinese (preliminary)

T has two different F_0 points: t_i and t_j, where t_i precedes t_j.
Let LZ (t_i) be the Logarithmic Z-score value of t_i.
Let LZ (t_j) be the Logarithmic Z-score value of t_j.
$|LZ\ (t_i)| - |LZ\ (t_i)| = d$.
If $|d| > 0.5$, T is a contour tone.

(15) Conversion of LZ score to the five-point scale (see text)

LZ-score	5-point scale	3-height scale
2	5	Extra High
1.5	4	High
1	$^{(!)}4$	High
0.5	3	Mid
0	$^{(!)}3$	Mid
-0.5	2	Low
-1	$^{(!)}2$	Low
-1.5	1	Extra Low

The two extreme points (5 and 1 or Extra High and Low) are "non-contrastive." For example, 52 and 41, can be both categorized as a "high falling tone." In other words, as we will see, 5 and 4 are in actuality not phonologically contrastive levels (as well as $^{(!)}2$ and 1). Second, the range of 4, 3 and 2 is within 1 in LZ, corresponding to High, Mid and Low, respectively. As was defined in (14), however, the threshold value of the F_0 fluctuation for a level tone is less than 0.5 in LZ. So it is possible that a contour tone may be positioned between 0 and -1 and has a Δ_{LZ} of 0.7, for example. The contour tones with a minor pitch excursion are represented with a raised exclamation mark on the second digit, e.g. $2^{!}2$.

There seems to be a three-way contrast of duration for the citation tones. Checked tones (Tone 7/8) are the shortest, which is expected because checked syllables are usually the shortest syllable type in Sino-Tibetan languages. However, it is striking to see that Tone 1/2 have only around half of the length of the other long tones (i.e. Tone 3/4/5/6). In particular, Tone 1 (551) has the greatest pitch excursion among the citation tones but its duration nevertheless is the shortest. Likewise, Tone 2 (231) is considerably shorter than the other L-register tones.

Tone	Pitch value	Duration: Min~Max	Contour shape
Tone 3/4/5/6	334/223/31/221	366.5 ms~430 ms	Rise/Gentle Fall
Tone 1/2	551/231	168.5ms~239.3ms	Steep Fall
Tone 7/8	55/23	89.2 ms~127.1 ms	Level/Rise

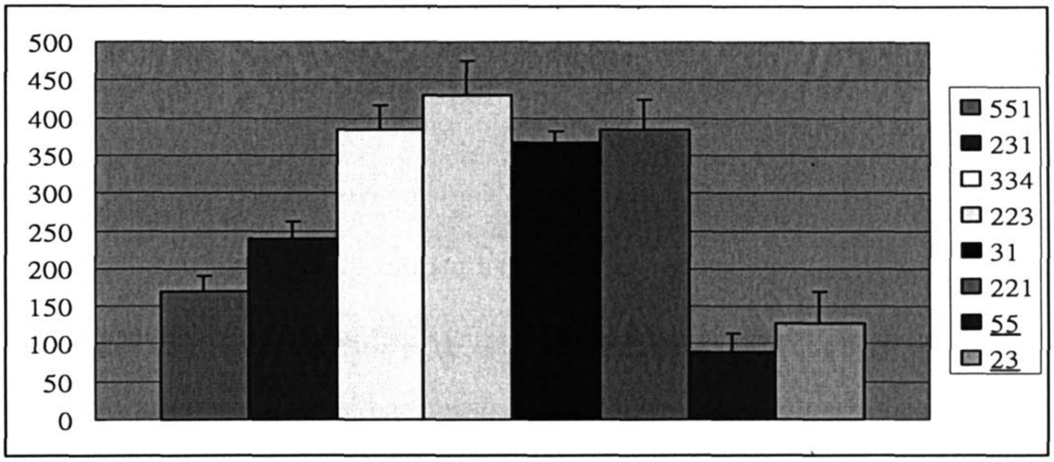


Figure 4-6 Normalized duration (in ms) of citation tones in Shaoxing Chinese

Finally, it appears that there are four falling tones in this inventory, namely, Tone 1 (551), Tone 2 (231), Tone 5 (31) and Tone 6 (221). This inventory is not predicted by any binary approach to tonal features (Recall the "too many contour tones" problem in §1.3.2). One common response is to deny that tones with a minor pitch excursion, e.g. Tone 6 (221), are phonologically falling tones. As proposed in the beginning of this section, the Δ_{LZ} of Tone 6 is still greater than 0.5. In addition, as we will see, Tone 6 (221) behaves as a contour tone in disyllabic tone sandhi. Recall the discussion at the outset. Falling tone must precede a contour tone in (3)a. We will see that this phonotactic restriction is attested in Tone 6 (221). Unless these two pieces of evidence are proven to be incorrect or misinterpreted, Tone 6 (221) is treated as a falling tone. That being the case, the tonal inventory of SX raises a non-trivial challenge to most of the models of tonal representation.

Above are description and discussion of the phonetic traits of citation tones. Before we move on to discuss what these properties tell us about the SX tonology, we describe the tonal alternations in disyllabic tone sandhi.

4.3 Acoustic Properties of Disyllabic Tone Sandhi

This section provides a phonetic description of SX disyllabic tone sandhi. Like many other varieties of Wu Chinese, two types of tone sandhi are distinguished in SX, namely, Initial-dominant and final-dominant sandhi. In Northern Wu Chinese to which SX belongs, the initial-dominant sandhi can be used in words and phrases of any syntactic structure (hence was called "general sandhi rules" in the descriptive literature), while final-dominant sandhi is used for syntactic configurations such as Subject-Predicate and Verb-Complement (hence was called "restricted sandhi rules"). The data reported here are of the initial-dominant sandhi.[8]

4.3.1 Tone 1/Tone 2-initial Patterns

Let us first observe disyllabic words for which Tone 1 (551) and Tone 2 (231) appear in initial position. The table in (16) summarizes all of the attested six (6) patterns. In the UR column (Underlying Representation), citation forms of both tones are given, followed by the SR column (Surface Representation). The initial tones and the final tones are separated by a hyphen. Parenthesized tones denote two specific combinations that can be collapsed, e.g. 551-(551/334) means that 551-551 and 551-334 have identical surface forms, 33-42. Mean duration for each tone is given and the standard deviation (SD) is given in parentheses. Finally, N means number

---

[8] All of the disyllabic words are Noun-Noun compounds or lexicalized Adjective-Noun phrases. Also, the experimental procedures are same as described in fn. 4. See Ping (2001: 101-102) for the wordlist.

of tokens.

As for the pitch tracks, data are arranged in the following format. The normalized and LZ normalized F_0 on the left hand side denote the initial tone, followed by a blank region. The blank region stands for a hypothetical intervocalic consonant and does *not* represent the actual C-interlude. The F_0 curves on the right hand side represent the tones on the second syllable.

Finally, the averaged absolute durations of the initial tone, the intervocalic consonant and the second tone are provided in separate histograms.

(16) Tone 1 (551) and Tone 2 (231) in initial position

Pattern	UR	SR	Mean duration (SD): T_1/T_2	N
A	551-(551/231)	33-42	198.3(33.9)/135.2(27.6)	15
B	551-(334/223/31/221)	33-44	250.2(40.2)/251.2(42.9)	18
C	551-(55/23)	33-44	259.9(47.3)/81.1(25.4)	16
D	231-(551/231)	22-42	247.4(56.5)/161.1(26.1)	17
E	231-(334/223/31/221)	22-33	254.2(29.5)/251.0(35.3)	20
F	231-(55/23)	22-44	257.5(43.1)/78.5(47.1)	13

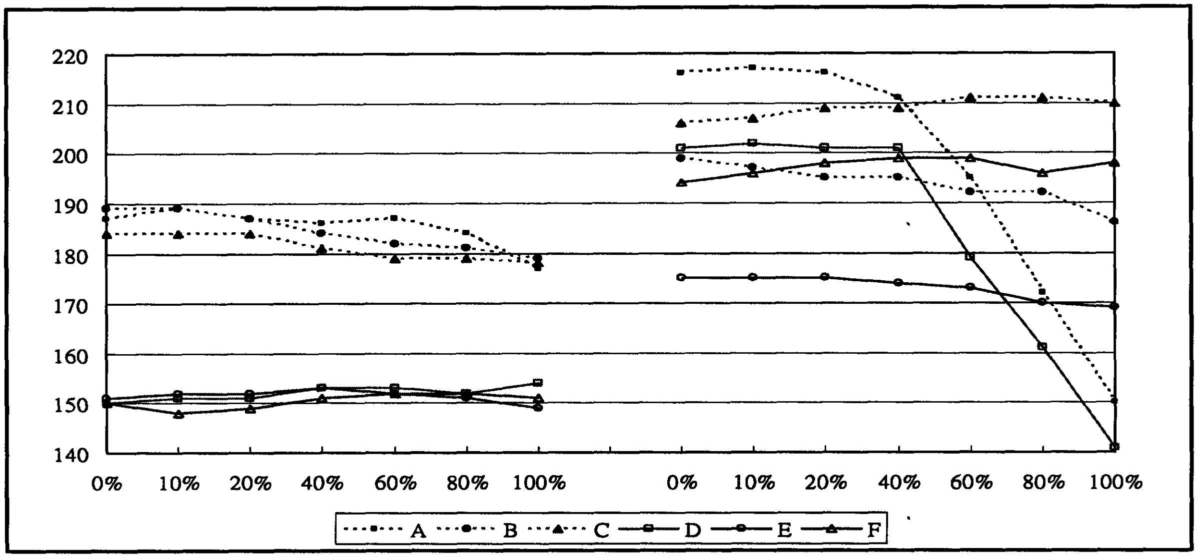


Figure 4-7 Normalized F_0 (in Hz) for Patterns A, B, C, D, E, and F

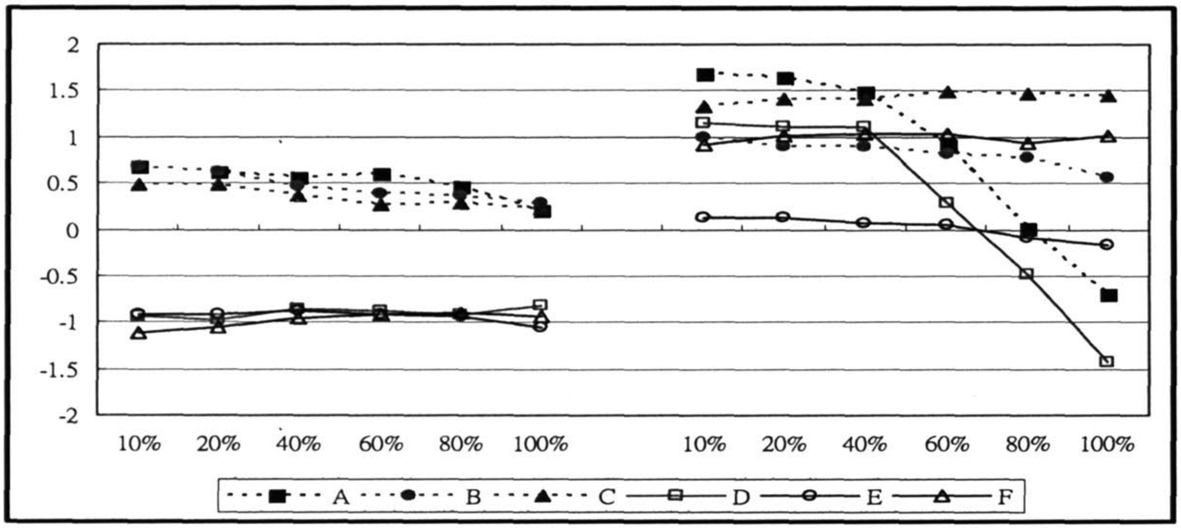


Figure 4-8 LZ normalized F_0 for Patterns A, B, C, D, E, and F

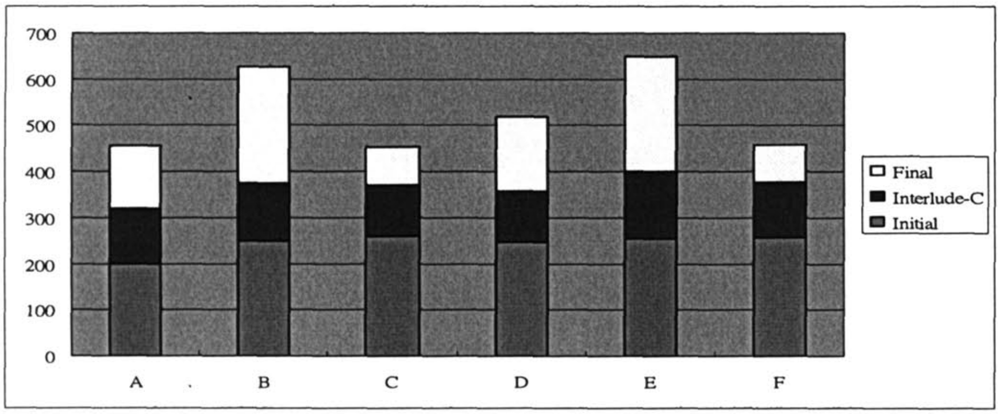


Figure 4-9 Stacked duration (in ms) for Patterns A, B, C, D, E, and F

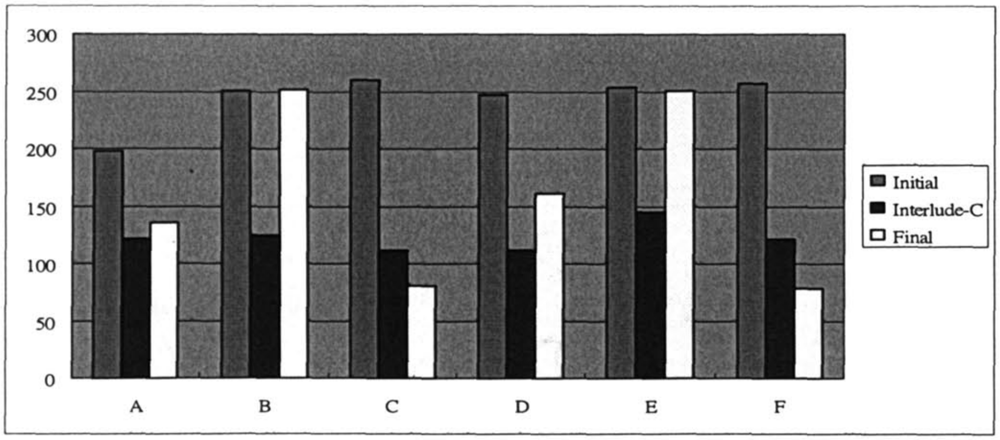


Figure 4-10 Clustered duration (in ms) for Patterns A, B, C, D, E, and F

There are several things to note about the preceding F_0 tracings. First, there is a robust H-/L-register distinction in initial position. By contrast, the register contrast is neutralized in non-initial position. As seen, the contours on the right hand side, i.e. the final tones, crowd in the upper part of the F_0 range. The LZ values of the non-initial tones are mostly greater than 0, whereas breathy voice-affected tones always start at around or below -0.5, suggesting that murmur is absent in non-initial tones. Therefore, it appears that the modal/non-modal phonation contrast is only preserved in initial position. Also, initial tones are preserved and non-initial tones undergo partial neutralization. So we may conjecture that the (non-)neutralization of tones seems to be crucially correlated with the presence/absence of breathy voice.

Second, regarding the contour shape, it is evident that Tone 1 (551) and Tone 2 (231) undergo complete contour reduction: word-initially, 551 becomes 33 and 231 changes to 22. The corresponding F_0 curves are flattened out, if compared with those in isolation (see also Figure 4-1 and Figure 4-2).

Third, all of the tones, checked or non-checked, are simplified to level tones in final position. Interestingly enough, Tone 1 and Tone 2 instead remain falling in this context. To this end, it should be fair to say that, in addition to their comparatively short duration in citation, it is obvious that Tone 1/2 do have some special properties that need to be explained.

Finally, with respect to duration, the initial tones are not always longer than the final tones. In Patterns A, C, D and F, the rime duration of the first syllable is indeed longer (approximately from 40% vs. 30% to 60% vs. 20%, including the interval). However, this observation does not hold for Patterns B and F, whereby the rime duration is basically equivalent (approximately 40% vs. 40%, again including the interval). In Patterns B and F, the second syllables carry Tone 3 (334), Tone 4 (223), Tone 5 (31) and Tone 6 (221). These are of the longest category in citation tones (cf. Figure 4-6). On the other hand, Tone 1 (551) and Tone 2 (231) belong to the second long category while the checked tones are the shortest in duration. The present data further show the "Tone 1/2 vs. Tone 3/4/5/6" difference.

4.3.2 Tone 3/Tone 4-initial Patterns

Let us now look at the tone patterns where rising tones (Tone 3/4) occupy the first syllable. All of the four (4) attested tone sandhi patterns are given in (17). As we expect, phonation registers remain distinct domain-initially and are neutralized in final position.

Regarding phonetic realization on the initial syllable, it is obvious that rising tones are not flattened out on the first syllable: the underlying contours are retained in sandhi initial position. Tone 3/4's Δ_{LZ} both exceed the proposed threshold value 0.5 in LZ. Remarkably, the pitch excursion of the L-register rising tone, Tone 4 (223), becomes greater to a significant extent in this context (i.e. 223 → 25).

With respect to the final tones, one of the striking facts is that all long non-checked tones are neutralized to a high falling tone (51). More surprisingly, checked tones are not exceptional, either. As illustrated in Patterns H and J, it is clear that Tone 7/8 both change to a high-to-mid falling checked tone (transcribed as $\underline{53}$).

Finally, unlike the patterns in the previous section, the initial tones are always longer than the final tones here, indicating that there is probably no canonical duration for all disyllabic combinations. Instead, the overall duration is largely conditioned by the initial tone. That is, if the initial tone is longer, then the final tone is longer.

(17) Tone 3 (334) and Tone 4 (223) in initial position

Pattern	UR	SR	Mean duration (SD): T_1/T_2	N
G	334-(551/231/334/223/31/221)	45-51	201.2 (26.8)/ 146.3 (24.0)	17
H	334-(55/23)	45-53	197.6 (55.5)/ 89.0 (34.2)	6
I	223-(551/231/334/223/31/221)	25-51	231.4 (35.4)/ 147.9 (26.6)	34
J	223-(55/23)	25-53	242.3 (27.1)/ 80.7 (15.5)	11

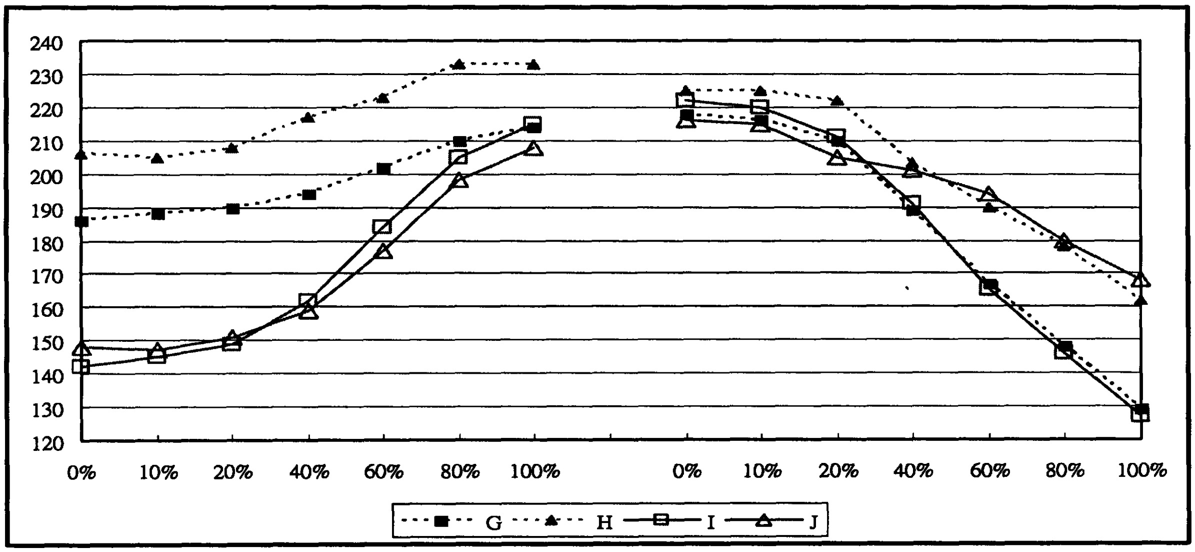


Figure 4-11 Normalized F_0 (in Hz) for Patterns G, H, I, and J

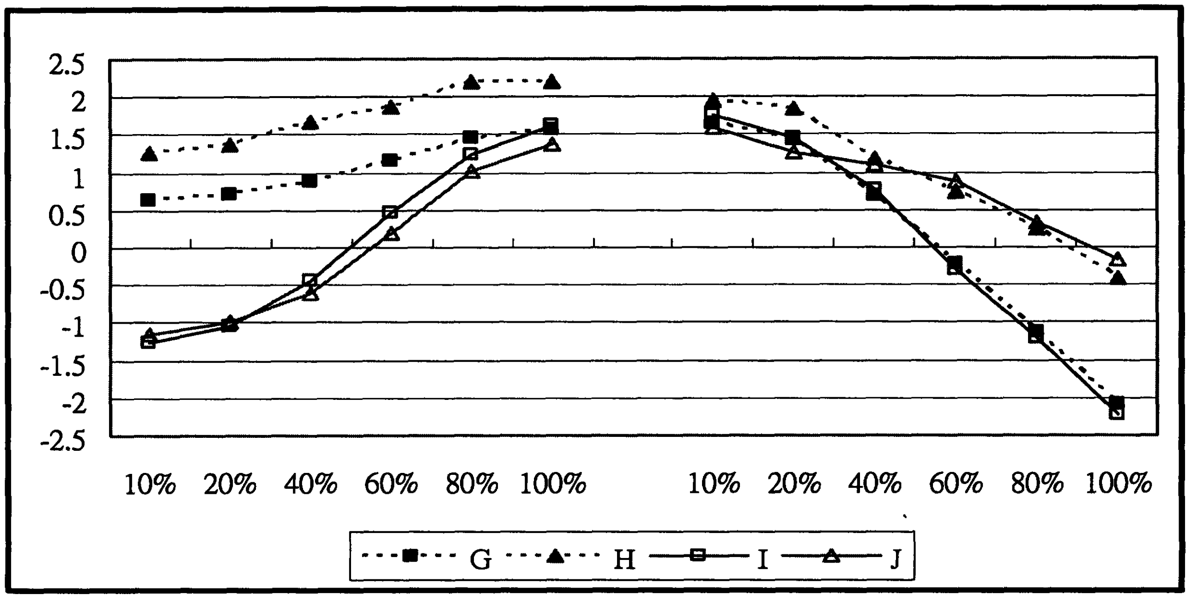


Figure 4-12 LZ Normalized F_0 for Patterns G, H, I, and J

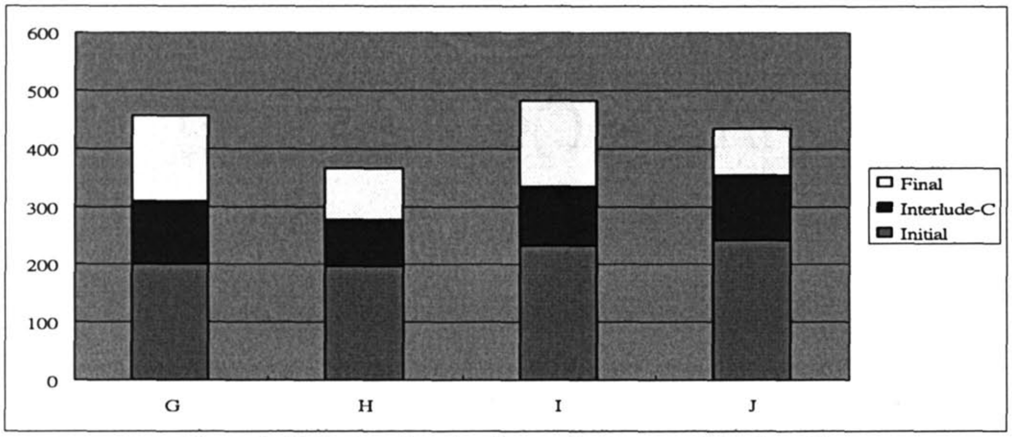


Figure 4-13 Stacked duration (in ms) for Patterns G, H, I, and J

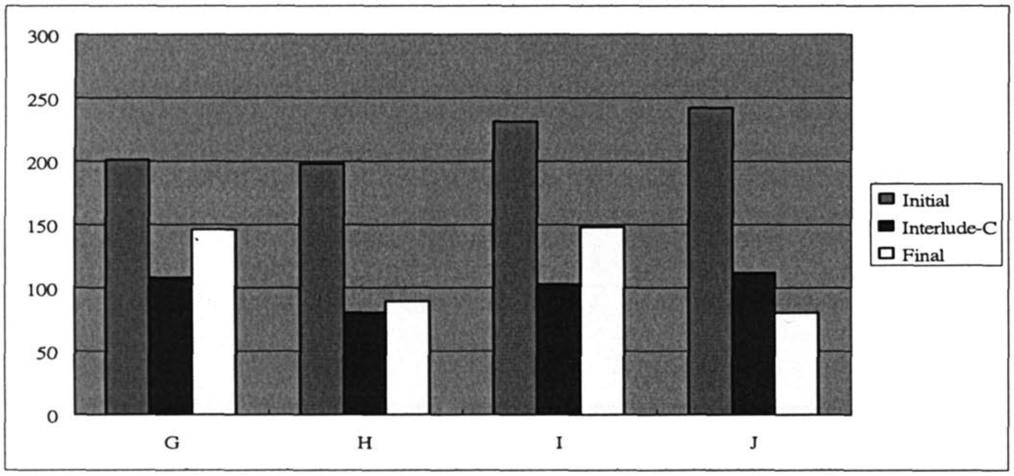


Figure 4-14 Clustered duration (in ms) for Patterns G, H, I, and J

4.3.3 Tone 5/Tone 6-initial Patterns

Tone 5 (31) and Tone 6 (221) both have comparatively minor pitch excursions in isolation. When Tone 5 occurs in initial position of a disyllable, this tone undergoes partial contour reduction (i.e. 31 → 32). The LZ-score differences between 10% and 100% are 1.58 (or 31 Hz) in citation and 1.01 (or 24 Hz) in initial sandhi position. By contrast, it looks as if not all of the contours of Tone 6 (221) are consistently falling. In Pattern N, its sandhi form has a small fall (LZ difference is still greater than 0.5), transcribed as $2^{!}2$. However, the sandhi forms of Tone 6 seem to be completely flattened in Patterns O and P as the F_0 trajectories fluctuate within the range of 0.5 in LZ. Assuming that an underlying tone has uniform surface realizations in a given context, we suspect that the initial tones of Patterns O and P should be treated as contour tones as well. So it appears that the proposed threshold value for level tones in (14) may not be sufficient.

Of special interest is the fact that the final syllables are generally L-register. As we can see, the tonal onsets of the second syllable are mostly below 0 in LZ or 160 Hz. In our discussion on Patterns A-J, all of those final tones rise to the H-register, suggesting that the phonation contrast is not licensed word-medially. This phonation neutralization also takes place for the present cases, because in addition to Patterns A-J, it has been instrumentally confirmed in a fair amount of Wu Chinese dialects that the phonation contrast is neutralized in *non-initial* sandhi position (e.g. Shi 1983, Iwata *et al.* 1991, Cao and Maddieson 1992, Ren 1992; see also fn. 6 for the list). Therefore, I assume that the L-register tones in final sandhi position are produced with modal phonation.

With regard to duration, the overall duration of Tone 3/4-initial patterns is the longest among all of the disyllabic combinations. It can be seen from the following data section that Tone 3/4 are of the longest duration in sandhi initial position as well as in sandhi final position. Durational reduction is limited to checked tones.

(18) Tone 5 (31) and Tone 6 (221) in initial position

Pattern	UR	SR	Mean duration (SD): T_1/T_2	N
K	31-(551/231/31/221)	32-21	255.6 (37.7) / 291.7 (42.3)	19
L	31-(334/223)	32-23	280.8 (49.9) / 317.5 (20.0)	10
M	31-($\underline{55}$/$\underline{23}$)	32-$\underline{34}$	309.4 (29.0) / 92.0 (17.7)	14
N	221-(551/231/31/221)	2ˈ2-ˈ2ˈ1	297.0 (36.6) / 332.5 (43.3)	23
O	221-(334/223)	ˈ2ˈ2-23	294.4 (36.4) /310.4 (45.8)	17
P	221-($\underline{55}$/$\underline{23}$)	ˈ2ˈ2-$\underline{23}$	302.1 (55.1) /104.1 (36.8)	12

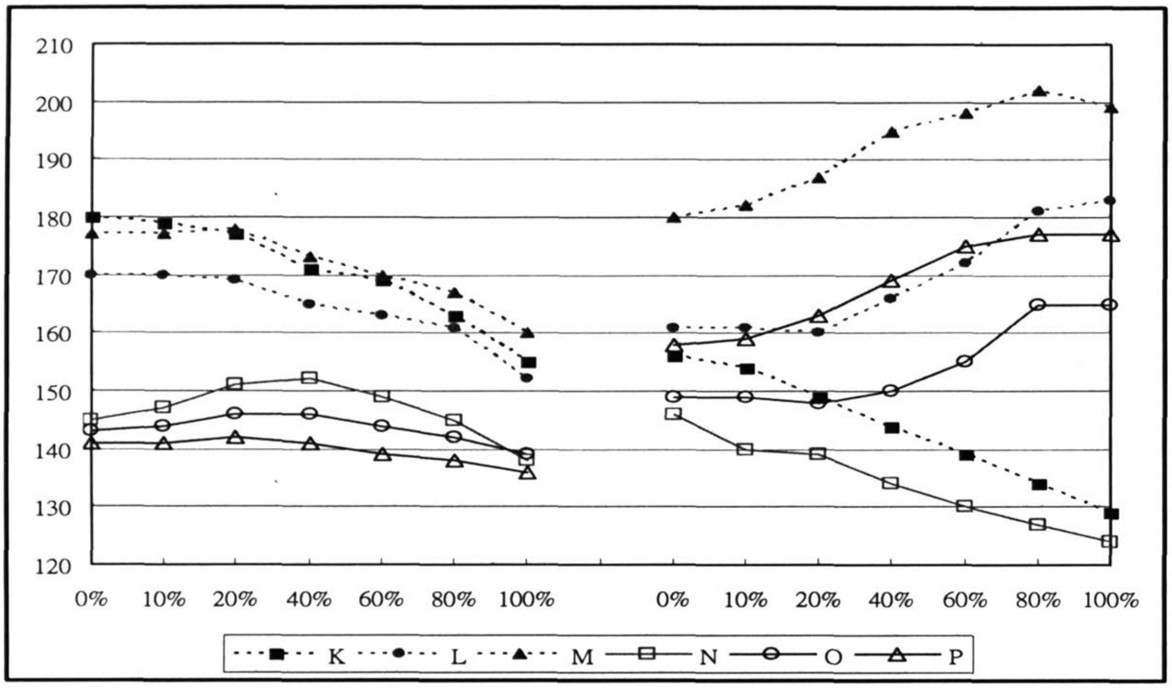


Figure 4-15 Normalized F_0 (in Hz) for Patterns K, L, N, O, and P

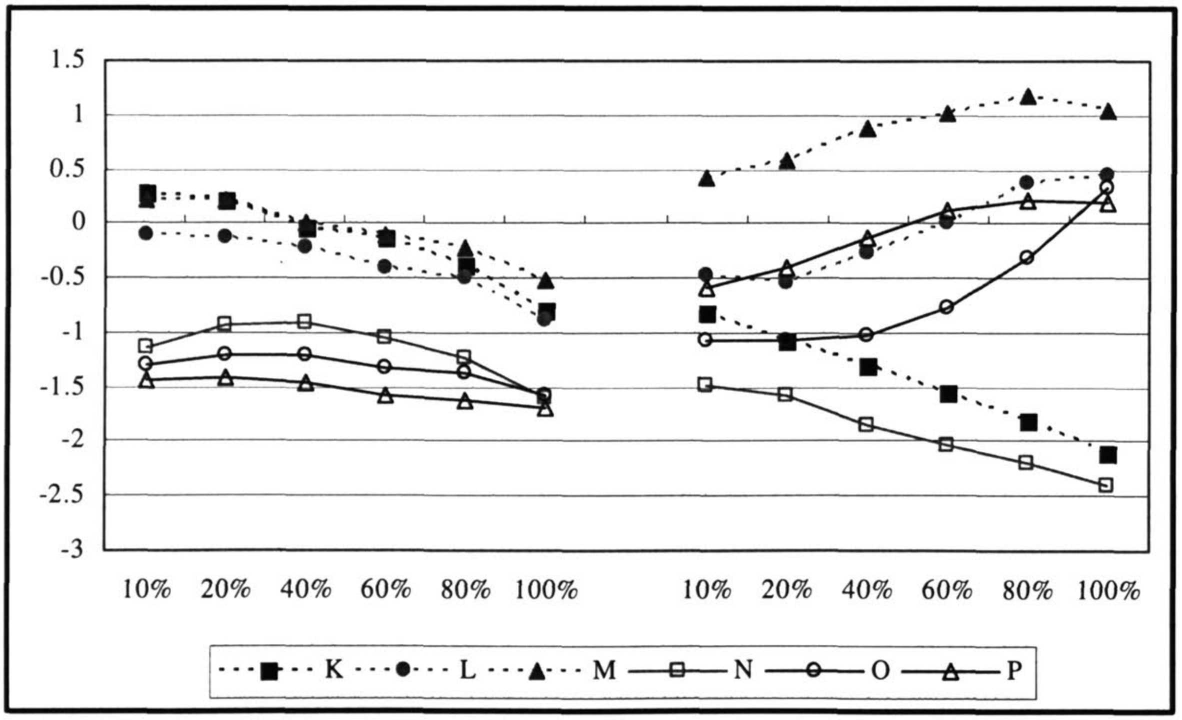


Figure 4-16 LZ Normalized F_0 for Pattern K, L, N, O, and P

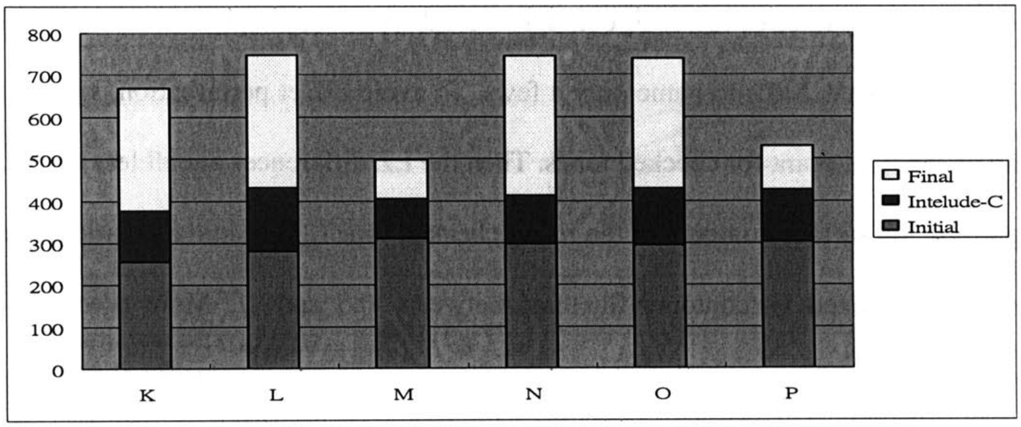


Figure 4-17 Stacked duration (in ms) for Patterns K, L, M, N, O, and P

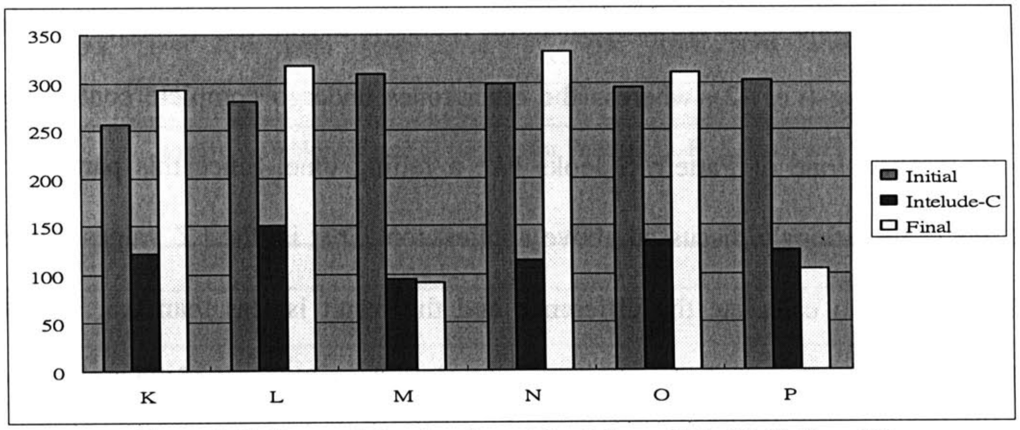


Figure 4-18 Clustered duration (in ms) for Patterns K, L, M, N, O, and P

4.3.4 Tone 7/Tone 8-intial Patterns

Disyllabic combinations with the checked tones on the first syllable are described in this section (19). The high level checked tone seems to become a falling tone in initial position (see Patterns Q, R and S). The LZ differences all exceed 0.5 if the 100% point is taken into consideration. Since these tones are terminated by a glottal stop, it is reasonable to assume that the F_0 lowering at the end is caused by vocal fry (or creaky voice). It has been reported that checked syllables in Sinitic languages are heavily glottalized (e.g. Taiwanese: Iwata *et al.* 1979; Cantonese: Iwata *et al.* 1981, Rose 2004; Shanghai Chinese: Zhu 1999) Even if the glottal stop coda drops in intersyllabic position, the creaky portion is retained. In addition to this, it has long been noted that creaky voice may raise or lower F_0 contours (Hombert 1978, Kingston 1985, Silerman 1997, Gordon and Ladefoged 2001, to name only a few). To avoid offset perturbation, we can take the 80% point as the final point for checked tones. Then the LZ differences are all less than 0.5.

Regarding the surface contours of the rising checked tone (23) in initial position, we can see that the LZ normalized F_0 contours fluctuate between -0.5 and -1. More precisely, the LZ differences between 10% and 80% of normalization duration are all smaller than 0.2. In other words, Tone 8 (23) is flattened out in initial sandhi position.

As far as the second tones are concerned, the recurrent asymmetry is observed: Tone 1 and Tone 2 remain falling (i.e. 42), whereas the other tones undergo complete contour reduction. Notably the second tone of Pattern S looks like a falling tone. Since this particular tone is checked, the same rationale discussed above applies, too. That is, the LZ values between 10% and 80% are used to calculate the difference and the result is less than 0.5. Therefore, the conclusion is that when checked tones are present in initial position, all of the tones except Tone 1/2 change to a level tone.

(19) Tone 7 (55) and Tone 8 (23) in initial position

Pattern	UR	SR	Duration (SD): T_1/T_2	N
Q	55-(551/231)	33-42	72.5 (15.3) / 157.1 (33.4)	19
R	55-(334/223/31/221)	33-33	68.3 (16.5) / 280.5 (33.6)	22
S	55-(55/23)	33-44	81.7 (19.7) / 90.4 (16.3)	13
T	23-(551/231)	22-42	74.9 (15.7) / 165.8 (33.3)	17
U	23-(334/223/31/221)	22-33	65.0 (17.1) / 254.3 (36.4)	18
V	23-(55/23)	22-44	83.1 (16.0) / 75.4 (12.4)	17

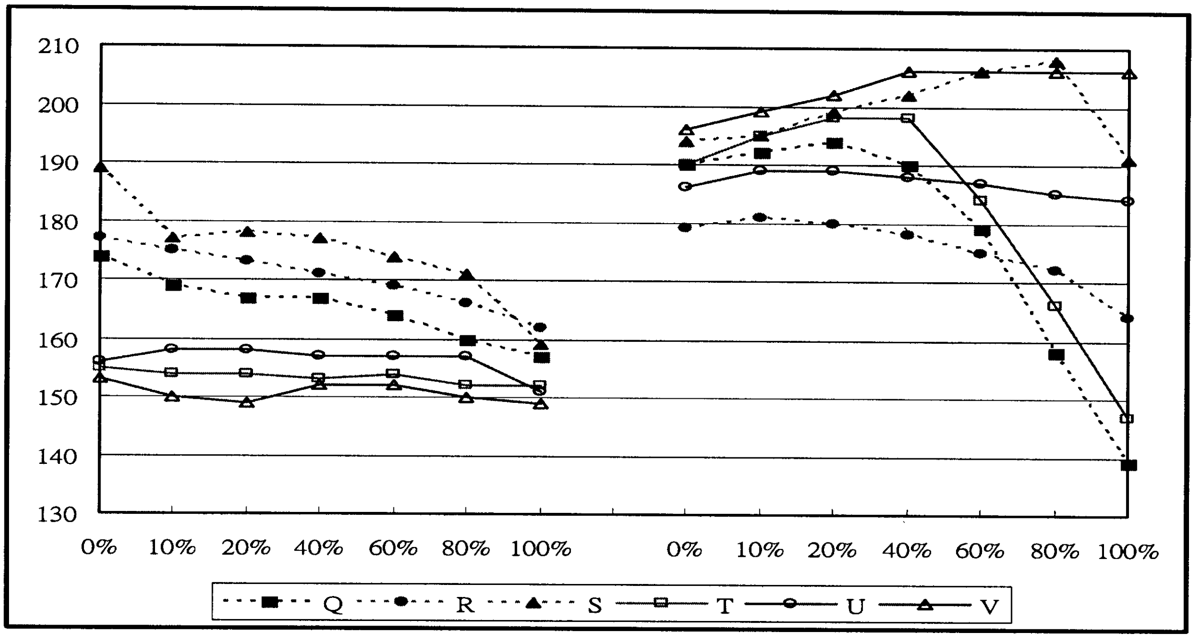


Figure 4-19 Normalized F_0 (in Hz) for Patterns Q, R, S, T, U, and V

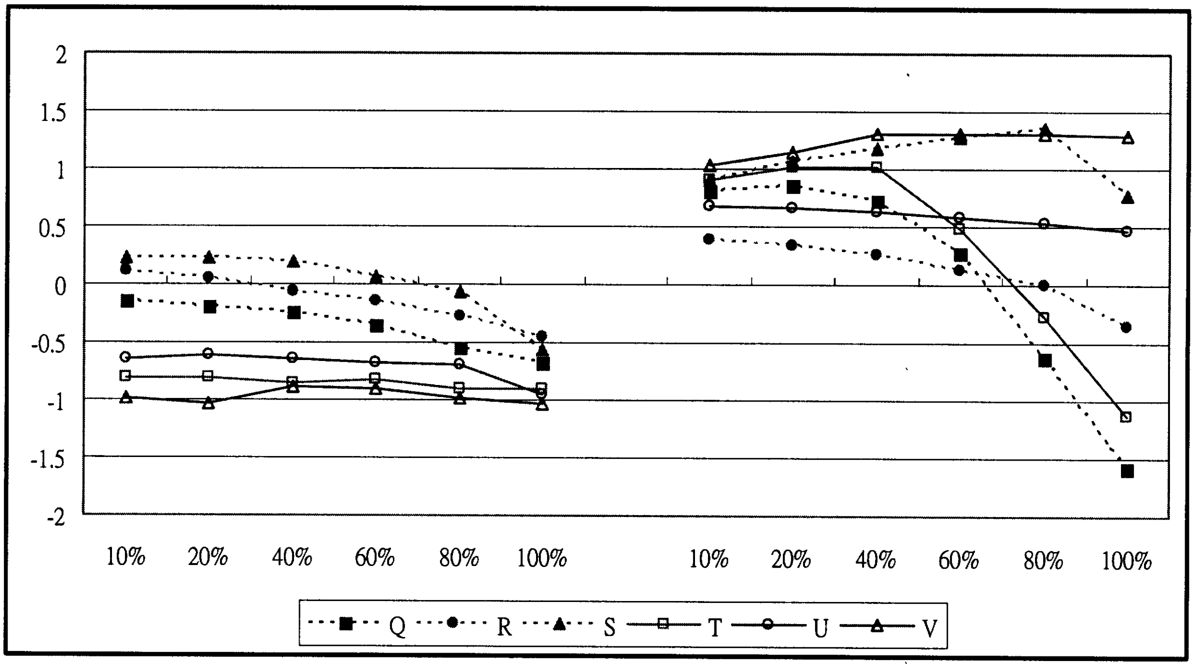


Figure 4-20 LZ Normalized F_0 for Patterns Q, R, S, T, U and V

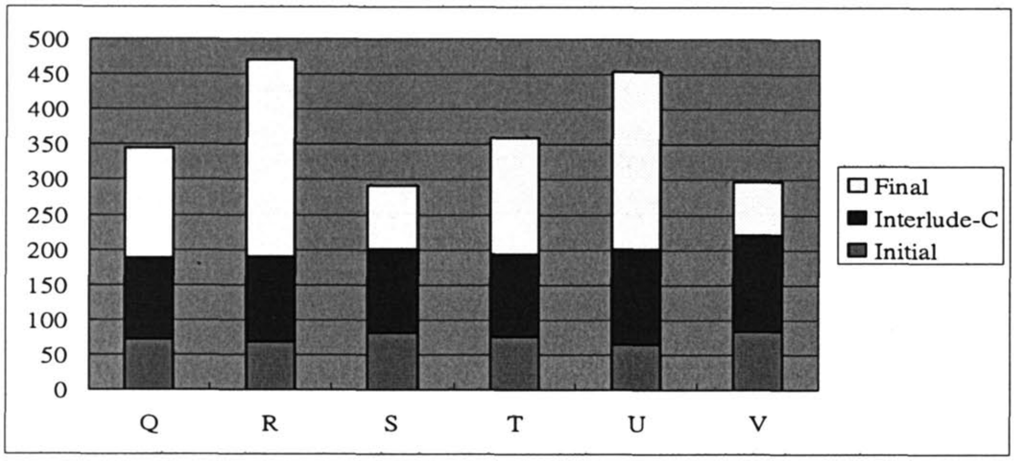


Figure 4-21 Stacked duration (in ms) for Patterns Q, R, S, T, U and V

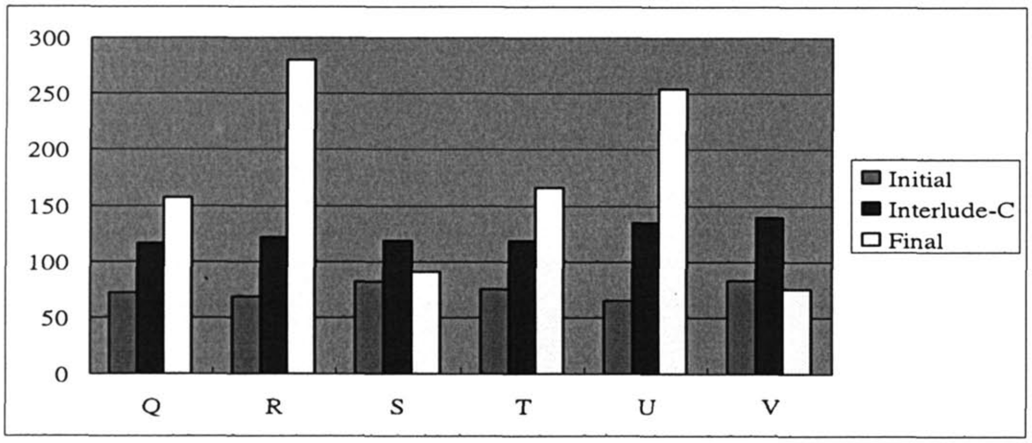


Figure 4-22 Stacked duration (in ms) for Patterns Q, R, S, T, U and V

Finally, initial checked tones are shorter than the final unchecked tones. As in §4.3.1, if Tone 1/2 are realized in final position, the length of their sandhi forms is considerably shorter than those of Tone 3/4/5/6, i.e. the "Tone 1/2 vs. Tone 3/4/5/6" distinction.

4.3.5 Duration of Disyllables

This section is a brief discussion of the duration of disyllabic words. The absolute and normalized duration of all twenty-two patterns are illustrated below. There are several things to note about the following histograms. First, the duration of the intervening consonants (i.e. the Interlude-C category) are largely constant. Second, the overall duration of a disyllabic word, including the initial rime, the interval and the final rime, is not fixed. We can say that there is no canonical duration for a disyllabic word. Interestingly enough, the general tendency is that, aside from the checked tones (i.e. Pattern C, F, H, J, M, P, Q, R, S, T, U and V), if the initial tone is longer, then the final tone is longer (i.e. Pattern B, E, K, L, N and O). Conversely, if the initial tone is shorter, then the final tone tends to be shorter (Pattern A, D, G and I). As mentioned earlier, this tendency may reflect that, aside from tonal contour, the initial tone also determines phonetic length of the following tone to a considerable extent.

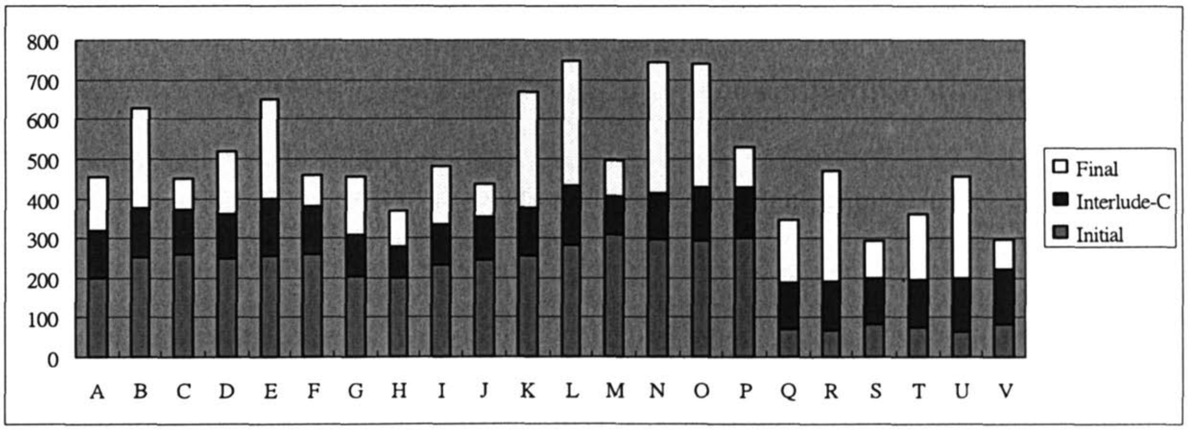


Figure 4-23 Stacked duration (in ms) for all disyllabic tone sandhi patterns

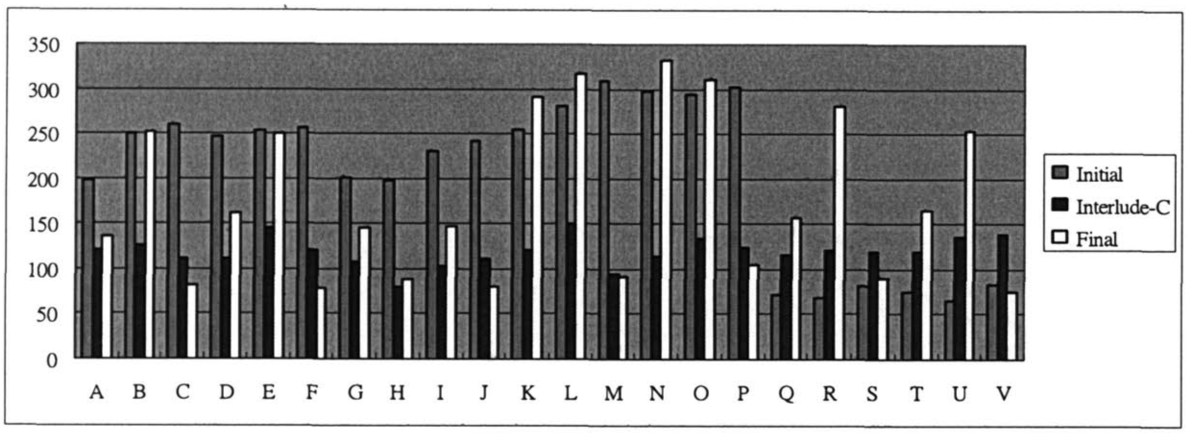


Figure 4-24 Clustered duration (in ms) for all disyllabic tone sandhi patterns

This completes the phonetic description of SX disyllabic tone sandhi. Bearing this range of tone sandhi data in mind, I turn in §4.4 to questions of what these phonetic data can tell about the phonology of SX tone sandhi.

4.4 Defining the Contour Tone

The core issue in this chapter hinges on the contour/level distinction (or the equal to/ non-equal to relation). As we have seen, SX has a comparatively large tone inventory. In addition to this, phonetic realizations of tones have quite intricate contextual variations. We then want to ask how we know whether Tone T is a level tone or a contour tone. As a widely accepted working definition, Maddieson (1978) suggests that the definition of a level tone is "one for which a level pitch is an acceptable variant." At least for Sinitic languages, it has never been clear how an "acceptable" variant should be determined for citation tones, let alone tones in context. This section thus serves as an attempt to deal with this issue in a more objective and quantitative fashion.

4.4.1 Positive skew in F_0 distribution in Shaoxing Chinese

As a first approximation, I have suggested that a difference less than 0.5 in LZ should be the threshold value for level tones in SX. But it turns out that this working definition does not work well in disyllabic tone sandhi. For example, some sandhi forms of Tone 6 (221) have a Δ_{LZ} less than the proposed threshold value 0.5. However, as we will see, phonological evidence indicates that those seemingly completely flattened tones behave like a contour tone.[9] So it appears that a fixed threshold value is not sufficient: the impression is that level tones in the higher part of the F_0 range allow a more pronounced fluctuation, whereas the acceptable variation for level tones in the lower part is narrower. As a reminder, we have discussed the *positive skew in F_0 distribution* in §1.3.1: the same F_0 interval (or tone spacing) in different pitch ranges does *not* necessarily stand for the same distance in production and perception. As evidenced in the data below,[10] tone levels in SX can be neatly characterized in this way. Note that subscripted letters denote a specific disyllabic tone sandhi pattern discussed in §4.3.

---

[9] It seems that some degree of circularity arises here. So some remarks are needed for clarification: by "phonological evidence," I mean the following phonotactic restriction: these seemingly completely flattened tones do not precede a non-contour tone (e.g. the initial tones in Pattern P and Q; see also Figure 4-25 and discussion thereof). I will argue that this restriction arises from the assumption outlined in §4.2, in particular, the schematic example in (3) and discussion that follows: if Tone T does not precede a level tone, then T is not supposed to be a level tone. Of course, this assumption is subject to justification. But the point is that aside from the observed alternations, I do not know of any means to characterize a linguistically relevant phenomenon solely with physical events. For example, a fall in 100 Hz may be contrastive in Language *P*, but not in Language *Q*. But how do we know whether 100 Hz is contrastive in the first place?

[10] Normalized F_0 values in Hz, but not in LZ, are used here. This choice is primarily because the anticipated asymmetry is more obvious as opposed to the logarithmically adjusted F_0.

(20) Difference of 10%-100% in Hz

	10%	20%	40%	60%	80%	100%	Δ(10%-100%)
22_D	151	151	153	153	152	154	3
22_E	152	152	153	152	151	149	3
22_F	150	148	149	151	152	151	1
33_A	187	189	187	186	184	177	10
33_B	189	187	184	182	181	179	10
33_C	184	184	181	179	179	178	6
55	202	204	206	210	213	214	12
33_Q	169	167	167	164	160	157	11
33_R	175	173	171	169	166	162	13
33_S	177	178	177	174	171	159	18

As seen, the F_0 fluctuations of 22, situated in the low range, are minimized to an extreme extent: the difference can be as small as 1 Hz throughout the entire pitch contour (Pattern F). In contrast, the mid level tone 33 allows a range from 10 Hz to 6 Hz. In addition to this, Tone 7 (55) is a checked tone, whose shorter duration is not supposed to afford a wider fluctuation. Even if the 80% point is taken as the final target (in order to avoid offset perturbation caused by the glottal stop coda), the difference remains essentially intact (i.e. 202 Hz (10%) – 213 Hz (100%) = 11 Hz). The above demonstration thus calls for a revision of the threshold value of an acceptable variant, i.e. a tone level, to which I turn in the following section.

4.4.2 Quantitative Criteria of Contouricity

The aim of this section is to quantify SX contour tones. It should be sufficient to use F_0 alone as the dimension to define a contour tone, although other dimensions such as duration may play a role, too. The fundamental assumption is that for a given tone, if the difference between its initial and final F_0 normalization points exceeds some threshold value, then that tone counts as a contour tone. For SX, I propose that the measurement points are designated as follows.

(21) Measurement points

Initial point	Modal: 10% / Breathy: 30%
Final point	Non-checked: 100 % / Checked: 80%

The reason why the initial point is the normalization point at 10% is to avoid onset perturbation. This is especially important in a language like SX with phonation contrasts at the tonal onset: as seen in the foregoing F_0 tracks, breathy voice depresses the initial F_0 contours and fades away in the middle point of the vowel (see also §4.2.1). So the 30% point, a point between the voicing onset and the mid point of a vowel, is chosen as the initial measurement point in case of murmured syllables. Likewise, for checked tones, it should be also appropriate to take the 80% point as the final measurement point because we need to minimize the influence of offset perturbation due to the glottal stop coda. The above assumptions allow us to quantitatively define the contour tones in SX as follows. Although the following definition is in essence *post hoc*, the contour/level distinction is nevertheless coherently manifested. In addition to this, we also have a plausible explanation why the threshold values should be so assigned: the (possible) logarithmic nature of the production and perception of F_0. So I will use this as a working hypothesis here, awaiting further research for a more motivated account.

(22) Definition of contour tone in Shaoxing Chinese

Let t_i and t_j be the measurement points for Tone T, where t_i precedes t_j.
Let LZ (t_i) be the Logarithmic Z-score value of t_i.
Let LZ (t_j)be the Logarithmic Z-score value of t_j.

a. T belongs to the high part of the F_0 range (above 1 in LZ)
$|LZ (t_i)| - |LZ (t_j)| = d$
If $|d| > 1$, then T is a contour tone.

b. T belongs to the middle part of the F_0 range (between 1~-0.5 in LZ)
$|LZ (t_i)| - |LZ (t_j)| = d$
If $|d| > 0.5$, then T is a contour tone.

b. T belongs to the lower part of the F_0 range (below -0.5 in LZ)
$|LZ (t_i)| - |LZ (t_j)| = d$
If $|d| > 0.2$, then T is a contour tone.

Put simply, if the LZ difference between two measurement points of Tone T exceeds 1/0.5/0.2 in the high/mid/low F_0 range, then T is a contour tone. Conversely, if the Δ_{LZ} among the normalized points in Tone T are all smaller than 1/0.5/0.2 in the high/mid/low part of the F_0 range, then T is a level tone. In order to (re)confirm the threshold values in different pitch ranges, let us observe the following F_0 tracings. In this figure, I represent the F_0 curves corresponding to 45 (whose citation tone is Tone 3 (334): G-1, H-1), 32 (whose citation tone is Tone 5 (31): K-1, L-1 and M-1) and $2^{!}2$ (whose citation tone is Tone 6 (221): N-1, O-1 and P-1). Note also that all of these tones are in initial sandhi position so N-1, for example, is used to indicate the initial tone of Pattern N.

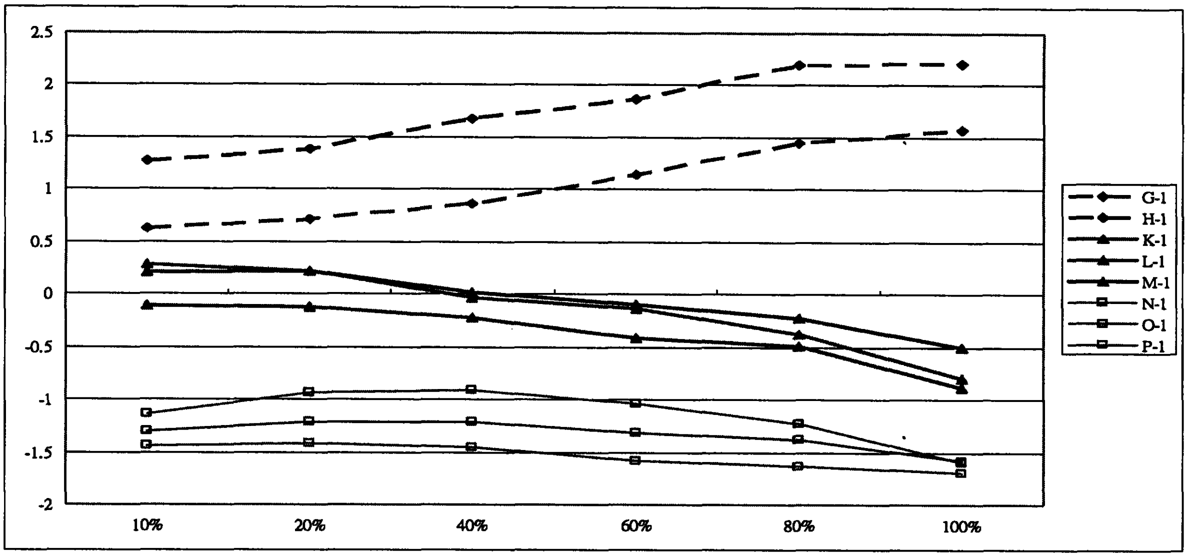


Figure 4-25 Initial sandhi contour tones in different ranges (NB: G-1 means the first tone in Pattern G and so on)

As we can see, the rise and fall is greater for the tones in the higher part (i.e. above -0.5 in LZ), whereas the fall is less pronounced in the low range (i.e. below -0.5). In particular, the curves corresponding to O-1 and P-1 (transcribed as 2ꜝ2) look very much like a level tone. Recall the discussion in (20): the initial (10%) and final (100%) target of 22 has a difference ranging from 1 to 3 Hz. For O-1 and P-1, the difference is 5 Hz (=144 Hz-139 Hz in O-1 and 141 Hz-136 Hz in P-1). Since 22 and 2ꜝ2 do not pattern alike in tone sandhi (cf. fn. 9), I take 3 Hz as the threshold value. When transformed into LZ, the corresponding value is 0.2. Therefore 0.2 is set up as the threshold value for a level tone in the lower part of the F_0 range. As discussed in the preceding section, a small F_0 interval in the lower part has the same distance as a greater F_0 interval in the higher part. The SX data may serve as an additional piece of evidence in support of the view according to which the production and perception of tone is based on a logarithmic relationship.

4.4.3 Summary

This section investigates several essential aspects of the tonal system of SX. Based on the phonetic description and discussion in §4.2 and §4.3, the key to the core issue in this chapter, the contour/level distinction (i.e. (non-)contouricity), has been defined in a quantitative (albeit *post hoc*) manner. Let us now move on to the argument for the core issue in this chapter, the non-equal to relation in relational correspondence.

4.5 (Non-)Contouricity Agreement as (Non-)Equal to

This section provides a formal analysis of the array of SX tone sandhi facts. I carry out my analysis under the assumption of the relational correspondence-based model. I begin in §4.5.1 with an overview of analytical issues. Given that the data are complex in their interactions, I break the discussion into three main sections. In §4.5.2, I address the core issue in this chapter: the implementation of the (non-)equal to relations in (non-)contouricity agreement I will argue that the non-contouricity agreement phenomenon is best accounted for through relational correspondence. Section 4.5.3 provides an analysis of phonetic realization of tones in initial position. In §4.5.5, the anomalous behaviors of Tone 1 and Tone 2 are examined. Finally, I deal with conditions on tonal register (dis)harmony §4.5.4.

4.5.1 Two Generalizations

Recall that the examples we have briefly discussed in §4.2. The phenomenon in question is such that a contour tone must be followed by a contour tone and a level tone must precede a level tone. To see if SX disyllabic tone sandhi patterns approximate the above statement, consider the tabulated data below.

(23) Disyllabic tone sandhi in SX

T2 / T1	551	231	334	223	31	221	$\underline{55}$	$\underline{23}$
551	*33-42*	*33-42*	33-44	33-44	33-44	33-44	33-$\underline{44}$	33-$\underline{44}$
231	*22-42*	*22-42*	22-33	22-33	22-33	22-33	22-$\underline{44}$	22-$\underline{44}$
334	45-51	45-51	45-51	45-51	45-51	45-51	45-$\underline{53}$	45-$\underline{53}$
223	25-51	25-51	25-51	25-51	25-51	25-51	25-$\underline{53}$	25-$\underline{53}$
31	32-21	32-21	32-23	32-23	32-21	32-21	32-$\underline{34}$	32-$\underline{34}$
221	2ꜝ2-ꜝ2ꜝ1	2ꜝ2-ꜝ2ꜝ2	2ꜝ2-23	2ꜝ2-23	2ꜝ2-ꜝ2ꜝ1	2ꜝ2-ꜝ2ꜝ1	2ꜝ2-$\underline{23}$	2ꜝ2-$\underline{23}$
$\underline{55}$	*$\underline{33}$-42*	*$\underline{33}$-42*	$\underline{33}$-33	$\underline{33}$-33	$\underline{33}$-33	$\underline{33}$-33	$\underline{33}$-44	$\underline{33}$-44
$\underline{23}$	*$\underline{22}$-42*	*$\underline{22}$-42*	$\underline{22}$-33	$\underline{22}$-33	$\underline{22}$-33	$\underline{22}$-33	$\underline{22}$-44	$\underline{22}$-44

(Note that the raised exclamation mark on the second digit, e.g. 2ꜝ2, denote a minor fall. Notice also that level tones are only those with the identical juxtaposed digits. See also (15) for the conversion table from LZ to the five-point scale.)

These data reflect two important generalizations. First, we can see that if the initial tone is rising, then the following tone must be falling. If the initial tone is falling, then the following tone can be either falling or rising. If the initial tone is level, then the following tone becomes level, except for the case of Tone 1 (551) and Tone 2 (231) on the final syllable (Recall the Tone 1/2 vs. Tone 3/4/5/6 distinction, *passim*).

(24) If T_1 is [αCONTOUR], then T_2 is [αCONTOUR].

a. If T_1 is a level tone, then T_2 must be a level tone.
b. If T_1 is a contour tone, then T_2 must be a contour tone.
c. Tone 1 (551) and Tone 2 (231) are 'exceptions' to this generalization.

Secondly, the (non-)contouricity of the initial tone conditions the tonal register of the following tone in a systematic way. If the initial tone is a level tone, the final tone is always H-registered. However, if the initial tone is a contour tone, the tonal register of the second tone agrees with the initial tone. Notice further that since the modal/murmur contrast is neutralized word-medially (cf.

§4.3.1), register here refers to H- and L- register (i.e. "higher" vs. "lower" tones in modal phonation). The second generalization is stated as follows.

(25) Tonal register (dis)harmony

d. If the initial tone is level tone, the following tone is H-register.
e. If the initial tone is contour tone, the following tone must agree with its precedent tone in *tonal* register.

The above is a brief sketch of the two most important generalizations in SX disyllabic tone sandhi. At this point, we have demonstrated and discussed essential aspects of the SX tone sandhi system, from both the empirical and theoretical perspectives. We now set out to validate the non-equal to relation by considering tonal alternation in final sandhi position. Our analysis begins with the first generalization: If T_1 is [αCONTOUR], then T_2 is [αCONTOUR].

4.5.2 Contouricity Agreement in Final Sandhi Position

For ease of discussion, a comprehensive list of the attested patterns is repeated as follows.

(26) a. If T_1 is a level tone, then T_2 must be a level tone.
b. If T_1 is a contour tone, then T_2 must be a contour tone (shaded cell).
c. Tone 1/2 are exceptions to the above generalizations.

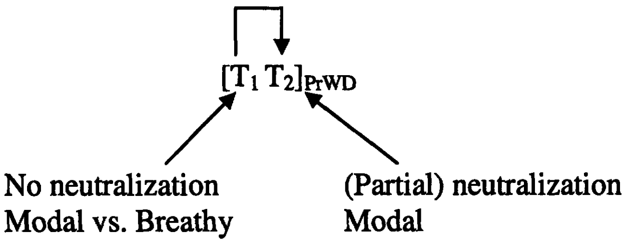
T_2 T_1	551	231	334	223	31	221	55	23
551	*33-42*	*33-42*	33-44	33-44	33-44	33-44	33-44	33-44
231	*22-42*	*22-42*	22-33	22-33	22-33	22-33	22-44	22-44
334	45-51	45-51	45-51	45-51	45-51	45-51	45-53	45-53
223	25-51	25-51	25-51	25-51	25-51	25-51	25-53	25-53
31	32-21	32-21	32-23	32-23	32-21	32-21	32-34	32-34
221	2ˈ2-ˈ2ˈ1	2ˈ2-ˈ2ˈ2	2ˈ2-23	2ˈ2-23	2ˈ2-ˈ2ˈ1	2ˈ2-ˈ2ˈ1	2ˈ2-23	2ˈ2-23
55	*33-42*	*33-42*	33-33	33-33	33-33	33-33	33-44	33-44
23	*22-42*	*22-42*	22-33	22-33	22-33	22-33	22-44	22-44

In terms of relational correspondence, the above generalization receives the following interpretation: the relation between the tone values associated with the two temporal spans in T_2 is identical to the relation between those in T_1. More precisely, (26)a amounts to saying that if the initial tone has the equal relation (x=y), then the final tone must have the equal to relation (x=y). Likewise, the description in (26)b is captured by the requirement that the initial tone and the final tone both exhibits the non-equal to relation.

Before we formalize the constraints for the agreement of the (non-)equal to relation, I would like to briefly summarize the driving force of tone sandhi in SX. We have learned from §5.2.1 that modal/non-modal phonation is contrastive only in word-initial position, while the phonation contrast is neutralized in non-initial position. This gives rises to tonal neutralization because in his perception experiments, Ren (1992) reports that the modal/murmur contrast serves as one of the main cues for tonal identification in Shanghai Chinese. Again, I assume that SX is no exception in this regard. The above discussion can be schematically summarized as follows.

(27) Identity between T_1 and T_2

AGREE(CONTOURICITY)



Once neutralized, tonal specification of the non-initial syllables may be heavily dependent on the initial syllables. Metaphorically speaking, we can say that the neutralized tone is free to copy everything from the non-neutralized tone, including *(non)-contouricity.* As a consequence, the similarity between the first tone and the second tone is greatly enhanced. This intuition can be captured by formulating the following AGREE constraints.

(28) AGREE-T_1T_2 ($x \neq y$)

Let T_1 be a tone contained within Rime R_i.
Let t_1 be a tone value contained within Rime R_i. Let S_1 be a temporal span associated with t_1
Let t_2 be a tone value contained within Rime R_i. Let S_2 be a temporal span associated with t_2.
S_1 precedes S_2.

Let T_2 a tone contained within Rime R_j.
Let t_3 be a tone value contained within Rime R_j. Let S_3 be a temporal span associated with t_3
Let t_4 be a tone value contained within Rime R_j. Let S_4 be a temporal span associated with t_4.
S_3 precedes S_4.

T_1 and T_2 are adjacent tones. T_1 precedes T_2.

Let $t_1 = x$, $t_2 = y$, $t_3 = a$, and $t_4 = b$.

If $x \neq y$, then $a \neq b$.

(29) AGREE-T_1T_2 ($x=y$)

Let T_1 be a tone contained within Rime R_i.
Let t_1 be a tone value contained within Rime R_i. Let S_1 be a temporal span associated with t_1
Let t_2 be a tone value contained within Rime R_i. Let S_2 be a temporal span associated with t_2.
S_1 precedes S_2.

Let T_2 a tone contained within Rime R_j.
Let t_3 be a tone value contained within Rime R_j. Let S_3 be a temporal span associated with t_3
Let t_4 be a tone value contained within Rime R_j. Let S_4 be a temporal span associated with t_4.
S_3 precedes S_4.

T_1 and T_2 are adjacent tones. T_1 precedes T_2.

Let $t_1 = x$, $t_2 = y$, $t_3 = a$, and $t_4 = b$.

If $x \neq y$, then $a \neq b$.

As discussed in §4.1, the above constraints motivate tone extension in an effort to avoid tone sequences that do not agree in the (non-)equal relations. In other words, tone sequences that are not similar in terms of (non-)contouricity will incur a fatal violation provided that the two AGREE constraints are top-ranked.

With these assumptions in mind, let us firstly turn in §4.5.2.1 to the issues arising from the contour tone-only sequences.

4.5.2.1 The Non-equal to relation at work: Contour tone-only sequences

This subsection tackles the contour tone-only sequences in disyllabic tone sandhi. As we can see from the shaded cells below, it is obvious that the initial contour tones must be followed by another contour tone, *regardless of* the underlying tonal specification on the final syllable. As mentioned in the preceding section, this is attributed to the tonal neutralization in non-initial position.

(30) Contour tone-only sequences (=shaded cells)

T₂ / T₁	551	231	334	223	31	221	55	23
551	*33-42*	*33-42*	33-44	33-44	33-44	33-44	33-44	33-44
231	22-42	22-42	22-33	22-33	22-33	22-33	22-44	22-44
334	45-51	45-51	45-51	45-51	45-51	45-51	45-53	45-53
223	25-51	25-51	25-51	25-51	25-51	25-51	25-53	25-53
31	32-21	32-21	32-23	32-23	32-21	32-21	32-34	32-34
221	2¦2-¦2¦1	2¦2-¦2¦2	2¦2-23	2¦2-23	2¦2-¦2¦1	2¦2-¦2¦1	2¦2-23	2¦2-23
55	*33-42*	*33-42*	33-33	33-33	33-33	33-33	33-44	33-44
23	*22-42*	*22-42*	22-33	22-33	22-33	22-33	22-44	22-44

The most straightforward evidence for contouricity agreement comes from the sandhi behaviors of the checked tones. In particular, Tone 7 (55) is no exception, changing to a high-to-mid falling tone 53 in this context. Recall that checked tones are terminated by a glottal stop and are of the shortest duration in SX (see also Figure 4-1). In other words, F_0 fluctuations should be minimized on a checked tone. With respect to contour markedness constraints, the following fixed ranking is widely accepted: *RISE » *FALL » *LEVEL. In addition to this, along the line of the duration-based approach to contour tone licensing (Gordon 1999, Zhang 2002), contour tone is claimed to be deterministically licensed by the duration of the sonorous rime. All else being equal (to which I will return shortly), it is impossible that a checked level should ever change to a checked falling tone because falling tones are assumed to be more complex than level tones. Especially on a durationally short checked syllable, the restriction on contour tone distribution should be more stringent. That is to say, a markedness reversal, i.e. a level contour changes to a falling one, is least expected on the shortest syllables.

The empirical fact, however, is that if preceded by a contour tone, the high level checked tone 55 has to change a high-to-mid falling checked tone 53. Again, it is evident that duration has nothing to do with the phenomenon in question: the mean duration of Tone 7 (55) is 89.2 ms in isolation and 89 ms in final sandhi position (Pattern H). A 0.2 ms difference is not meaningful in any sensible way. In addition, it is the initial syllable, but *not* the final syllable that is supposed to be stressed in a disyllabic compound because underlying tones are retained word-initially. That is to say, the final checked tone is stressless (hence undergoes neutralization or deletion according to most researchers). Instead, monosyllables in isolation are normally regarded as carrying stress in Sintic languages (e.g. Duanmu 1990). So stress, if any, has nothing to do with contour tone licensing, too: 53 appears an "unstressed" syllable whereas 55 occurs on a "stressed" syllable. All in all, the present data weaken the proposal according to which duration (or other phonetic correlates inducing length distinction) plays a crucial role in contour tone licensing, at least in SX.

As for the standard autosegmental approach, it should be obvious that employing unitary contour feature in AGREE(F), e.g. AGREE-T_1T_2[+RISING/+FALLING] makes wrong predictions because the surface sequences can be rise-fall or fall-fall. In particular, the *rise-fall* sequence is not expected under this approach. Using the mora is hopeless, too. Since a checked syllable is assumed to be monomoraic throughout, there is no apparent reason why a level high checked tone should become a high-to-mid one here. Finally, a boundary tone-based analysis simply fails to explain why a boundary L tone *only* docks onto the high level checked tone that follows a contour tone (which yields a falling checked tone), but nowhere else.

It should be now clear that existing approaches all fail to account for the contour tone-only sequence in a satisfactory fashion. The proposed analysis then proceeds as follows. Note that we will discuss phonetic realization of the initial syllable in §4.5.3.1. One representative example is analyzed in the following tableau. Notice again that only tonal alternations of the *final tone* are under consideration.

(31) 223-55 → 25-53: Rising-Checked level → Rising-Checked Falling

223-55	AGREE-T_1T_2 ($x{\neq}y$)	RELCORR(x=y)$_{\text{NUC}}$
☞ a. 25-53		*
b. 25-55	*!	

(NB: In §4.5.3, I will explain why 223 → 25 in initial sandhi position. So the initial tone is not analyzed here. In addition, I will explain in §4.5.5 why other possible candidates, e.g. 25-35 are eliminated.)

The RELCORR(x=y)$_{\text{NUC}}$ constraint is dominated so that the attested output form is correctly chosen. Candidate (b) loses out because the two tones do not agree in the non-equal to relation: 25 -55 is a contour ($x{\neq}y$)-level (x=y) tone sequence. By contrast, candidate (a) is selected as the winner because in this case both tones manifest the 'non-equal to' relation at surface: 25 -53 is a contour ($x{\neq}y$)-contour ($x{\neq}y$) sequence. So far, it should be clear that the contour tone-only sequences can be straightforwardly accounted for by employing the 'non-equal to' relation.

Furthermore, it is instructive to compare the optimal candidate (31)a with certain imaginable alternatives involving manipulations of steepness. First, not mentioned in the previous tableau is the potential candidate (25-)51. In this candidate, the level checked tone surfaces as a steep fall: 51. This steep fall does not surface because of undershoot because 51 is articulatorily impossible on a 89 ms long rime. Our argument is built on Xu & Sun's (2002) equation for the relation of minimum time of pitch falls as a function of pitch change size.

(32) $t = 89.6 + 8.7d$

where t is the amount of time (ms) it takes to complete a pitch shift and d is the size of pitch change in semitone (st). Take the 51 in Pattern H for example. Its pitch fall is 5.69 st (=225Hz – 162 Hz). According to (32), the minimum time to complete this size of pitch change is 140.24 ms, which is much longer than the reported length, 89 ms. So 51 loses out due to this physical articulatory impossibility. Likewise, I assume that the same is true of 52. That is, 89 ms is still not sufficient to complete a pitch shift of 52. Furthermore, the reason why 54 is not the actual output form is attributable to the fact that 54 is positioned in the higher part of the F_0 range. As we have discussed in §4.4.2, the threshold value for level tones is greater than 1 in LZ in this region. We can see from Figure 4-11 that the highest LZ value for the curves on the right hand side start at 2. So 54 roughly corresponds a fall from 2 to 1 in LZ. Accordingly, 54 still counts as a level tone. To this end, there should not be doubts that the optimal output can only be 53.

One final issue arises in the case of the non-checked final tones among contour tone-only sequences: as we have seen from (30), phonetic realizations of the non-checked tones after the rising tones, Tone 3 (334) and Tone 4 (223), are invariably a steep fall, 51. More interestingly, these steep falling tones appear on a comparatively short rime: 146.3 ms (SD=24; Pattern G) and 147.9 ms (SD=26.6; Pattern I). According to the equation in (32), the minimum time to complete a steep fall 51 is 167.2 ms (=89.6+8.7*8.92, based on Pattern G). So there is no problem with this size of pitch fall on an about 150 ms rime. The real challenge is why 53 are not chosen as the optimal surface realization.[11] If a more pronounced fall involves more effort from the laryngeal muscles, 51 is supposed to be less favorable than 53 in terms of minimization of articulatory effort.

However, I would like to point out that this popular view may not be as well-grounded as it appears to be. First, we have seen that high falling tones are significantly shorter than mid-falling tones in SX. Although a more comprehensive phonetic survey is definitely needed to test the robustness of this tendency, the SX data indicate that a more pronounced fall in pitch is not necessarily longer than a less pronounced fall in duration.

Second, more importantly, Hirose (1981) remarks that "the activation of muscle is achieved by asynchronous excitation of many different motor units, whereas at the time of relaxation all the units can stop their activity almost synchronously." Furthermore, Hallé (1994) EMG study of Mandarin Chinese tone production (taken together with Erickson's (1976) Thai data) report that "speakers with high-pitched voice can produce rapid high-to-low F_0 falls by simply relaxing F_0-raising activities." In contrast, "speakers with a lower-pitched voice additionally utilize an F_0-lowering device." His findings are instructive in that a high-falling tone may not be always more effortful than a mid- or low-falling tone. Instead, it is likely that tone production should be tuned into language-specific articulatory mechanisms. So it might well be the case that 51 is

---

[11] 52 may be undistinguishable from 51 in that tone levels 1 and 2 are in general non-contrastive (see (15)). As for 54, this candidate is treated as a level tone. See the preceding discussion on 54.

articulatorily equally or less effortful than 53. The preference of a steep fall over a smooth fall may be otherwise motivated, e.g. in terms of perceptual saliency.

To this end, the phenomena in question suggest that surface tone sequences must agree in the non-equal to relation. Other factors such as articulatory or perceptual principles should be taken in consideration as well. In the following section, I show that the other major type of restrictions on surface tone concatenations, namely, level tone-only sequences can be accounted for in the same fashion.

4.5.2.2 The Equal to relation at work: Level tone-only sequences

In this subsection, I demonstrate that the above analysis extends directly to level tone-only sequences. In the same vein, the data are explained by the requirement of which T_1 and T_2 must agree in the equal to relation.

(33) Level-only sequences (=shaded cell)

T_2 T_1	551	231	334	223	31	221	$\underline{55}$	$\underline{23}$
551	*33-42*	*33-42*	33-44	33-44	33-44	33-44	33-$\underline{44}$	33-$\underline{44}$
231	*22-42*	*22-42*	22-33	22-33	22-33	22-33	22-$\underline{44}$	22-$\underline{44}$
334	45-51	45-51	45-51	45-51	45-51	45-51	45-$\underline{53}$	45-$\underline{53}$
223	25-51	25-51	25-51	25-51	25-51	25-51	25-$\underline{53}$	25-$\underline{53}$
31	32-21	32-21	32-23	32-23	32-21	32-21	32-$\underline{34}$	32-$\underline{34}$
221	2ˈ2-ˈ2ˈ1	2ˈ2-ˈ2ˈ2	2ˈ2-23	2ˈ2-23	2ˈ2-ˈ2ˈ1	2ˈ2-ˈ2ˈ1	2ˈ2-$\underline{23}$	2ˈ2-$\underline{23}$
$\underline{55}$	*$\underline{33}$-42*	*$\underline{33}$-42*	$\underline{33}$-33	$\underline{33}$-33	$\underline{33}$-33	$\underline{33}$-33	$\underline{33}$-44	$\underline{33}$-44
$\underline{23}$	*$\underline{22}$-42*	*$\underline{22}$-42*	$\underline{22}$-33	$\underline{22}$-33	$\underline{22}$-33	$\underline{22}$-33	$\underline{22}$-44	$\underline{22}$-44

The analysis is essential the same as what we have discussed in the preceding section. The only difference is that now the constraint that is responsible for the switch from non-level to level is AGREE-T_1T_2 (x=y). For the present case, this constraint is active in ruling out a tone sequence where the initial tone is level but the final tone is not.

(34) 55-31→ 33-33: Checked level-Falling → Level-Level

55-31	Agree-T_1T_2 (x=y)	RelCorr(x>y)$_{\text{Nuc}}$
☞ a. 33-33		*
b. 33-31	*!	

This completes our discussion of the level/contour tone-only sequences. Under a relational correspondence-based approach, the switch from contour to non-contour or level to non-level can be neatly accommodated within the (non-)equal to relations.

The case of Tone 1/2 in contouricity agreement will be separately addressed in §4.5.4. I turn now to the analysis of the sandhi initial tones.

4.5.3 Phonetic Realization of Tones in Initial Sandhi Position

The use of the phrase "phonetic realization of tones in initial sandhi position" is to highlight the following observation: the underlying contours are mostly retained in this context. This is because the phonation contrast is licensed in word-initial sandhi position. For ease of discussion, the relevant phonetic data are summarized below.

(35) Phonetic realization of tones in initial sandhi position

		Citation		Sandhi	Description
a.	Tone 1/2:	551/231	→	33/22	*Sharp* Falling → Level
b.	Tone 3/4:	334/223	→	45/25	Rising → Rising
c.	Tone 5/6:	31/221	→	32/22[1]	*Smooth* Falling → Falling
d.	Tone 7/8:	55/23	→	33/22	Level/Rising → Level

The behaviors of the falling tones (Tone 1/2/5/6) are not uniform in this context. As seen, Tone 1 and Tone 2 (call them Type 1 falling tones) undergo contour leveling (35)a, whereas Tone 5 and Tone 6 (call them Type 2 falling tones) remain a less pronounced fall (e.g. 31 → 32) in initial position (35)c. On the other hand, the non-checked rising tones (Tone 3/4) remain a rise in pitch under the current circumstance. The checked rising tone (Tone 8) is reduced to a level tone presumably because of diminished rime duration. Finally, the high checked tone stays level but is lowered to the midrange (i.e. $\underline{55} \rightarrow \underline{33}$). In sum, although we can say that underlying tones are more or less faithfully realized in initial position, we should keep in mind that a satisfactory analysis of tonal realization in this context must account for the following analytical issues.

(36) Issues to be addressed for phonetic realization in initial sandhi position

a. Why Type 1 falling tones are flattened out but not Type 2 falling tones?
b. Why rising tones remain a rise, given that Type 2 falling tones are completely flattened?
c. What motivates pitch lowering of Tone 1/7?

Now we are ready to start our analysis. The first issue under examination concerns "contour preservation" in initial sandhi position.

4.5.3.1 Contour Preservation in Initial Sandhi Position

The underlying contours are preserved for the following tones in initial sandhi position: Tone 3 (334), Tone 4 (223), Tone 5 (31) and Tone 6 (221). Among them, Tone 3 (31) and Tone 4 (221) undergo partial contour reduction, surfacing as 32 and $2^{!}2$. Given time pressure, it is expected that these two tones are realized with minor pitch excursions in this context. Some relevant phonetic data are given below. Notice that the duration of the initial tone is taken from Pattern K for Tone 5 and Pattern N for Tone 6. These two sandhi tones both are of the shortest duration among Tone 5/6's word-initial sandhi forms (cf. (18)).

(37) F_0 and duration of Tone 5 and Tone 6 in citation and sandhi

			ΔF_0(10%-100%)	ΔLZ(10%-100%)	Duration
Tone 5	Citation	31	31 Hz	1.58	366.5 ms
	Initial	32	24 Hz	1.01	255.6 ms
Tone 6	Citation	221	16 Hz	0.88	383.7 ms
	Initial	22$^{!}$	9 Hz	0.65	297.9 ms

The slope (*m*) of citation Tone 5 is 0.08 (= 31 Hz / 366.5 ms)[12] and the slope of its sandhi form in initial is 0.09 (= 24 Hz / 255.6 ms). It should be fair to say that this is a nearly perfect match in steepness. Likewise, the slope of Tone 6 is 0.04 (= 16 Hz / 383.7 ms) and the slope of its word-initial correspondent in a disyllable is 0.03 (=9 Hz / 297.9 ms). Again, we can draw the conclusion that the steepness is barely changed in citation and sandhi forms. In sum, it should be evident that the present partial contour reduction is attributed to diminished duration under this circumstance.

On the other hand, rising tones (Tone 3/4) do not undergo contour reduction in initial sandhi position even if the rime duration is considerably diminished. Instead, their pitch excursions are both significantly augmented. Observe now the phonetic data below. Notice that the duration of the initial tone is taken from Pattern G for Tone 3 and Pattern H for Tone 4. Unlike the case of Tone 5/6, these two sandhi tones are of the longest duration among Tone 3/4's word-initial sandhi forms.

---

[12] Slope (*m*) is calculated according to the following equation.

$$m = \frac{F_0(t_i) - F_0(t_j)}{\mathrm{Time}(t_i) - \mathrm{Time}(t_j)}$$

where $t_{i/j}$ are targets of Tone T and t_j precedes t_i. See also chapter 5 for further discussion.

(38) F_0 and duration of Tone 3 and Tone 4 in citation and sandhi

			ΔF0(10%-100%)	*ΔLZ(10%-100%)*	Duration
Tone 3	Citation	334	13 Hz	0.62	384.1 ms
	Initial	45	24 Hz	0.86	201.2 ms
Tone 4	Citation	223	26 Hz	1.29	430.0 ms
	Initial	25	66 Hz	2.66	231.4 ms

In order to preserve the rise, it turns out that steeper slope in sandhi is inevitable here. The slope of Tone 3 (334) in citation is 0.03 (= 13 Hz/384.1 ms). If the steepness between citation and sandhi tones is required to be as close as possible, the rise in sandhi would be 6 Hz (= 201.2ms*0.03). According to the definition of contour tones (22), 6 Hz is within the threshold value for level tones in the higher part. In other words, Tone 3 is at risk of being flattened out. Consequently, in order to avoid contour leveling, the pitch range of Tone 3 has to be enlarged on a shorter syllable.

As far as Tone 4 (223) is concerned, as we can see in the following figure, its sandhi form seems to have an excessive pitch excursion (I-1 and J-1; note also that G-1 and H-1 are the sandhi forms of Tone 3 (334): 45). It appears that the curve corresponding to its citation form (Tone 6) is *not* at risk of neutralization with its surrounding tones (i.e. A-1, B-1, C-1, K-1, L-1 and M-1): the final targets are separated by at least 0.5 in LZ. Recall from (22) that 0.5 LZ is the threshold value in the midrange (1~-0.5).

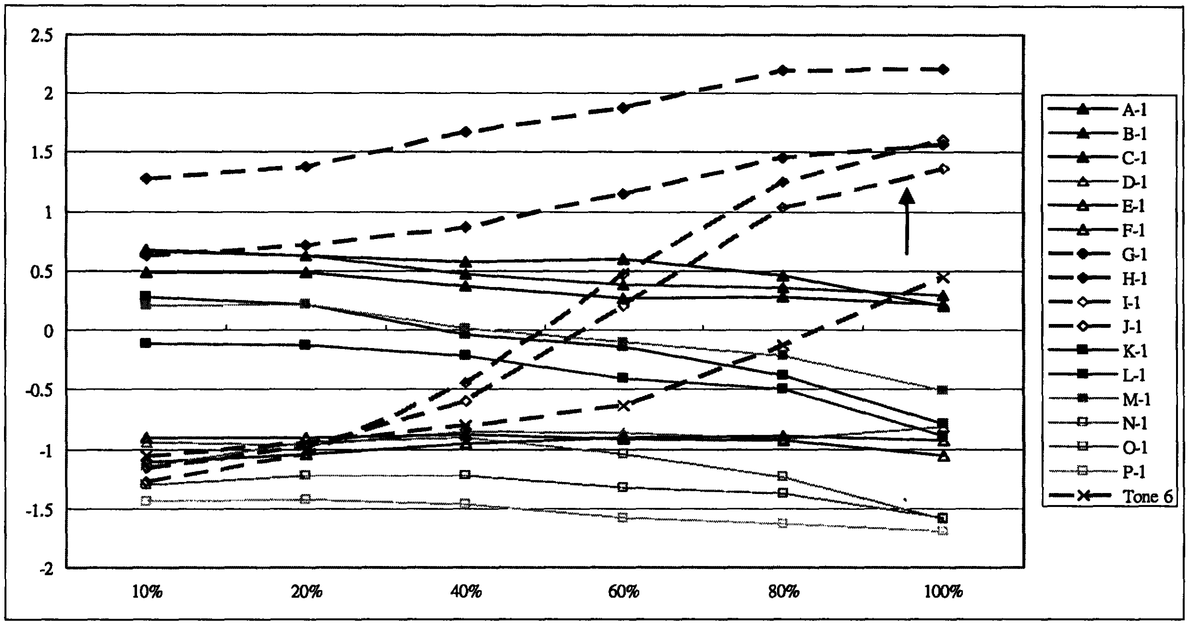


Figure 4-26 All of the non-checked tones in initial sandhi position (A-1 means the first tone in Pattern A, and so on)

How do we explain the overkill in pitch excursion? There seems no apparent reason as to why (and how) 223 should change to 25. I briefly mentioned in §4.3.2 that the long non-checked tones could be allotones (in a pair-wise fashion): the initial F_0 alternations are conditioned by the phonation contrast: modal phonation goes with high and murmur low. Let us suppose that Tone 3 (334) and Tone 4 (223) are allotones. Then it is reasonable to say that the paired tones should be "as similar as possible." The allotonic identity can be regulated by the following input-output correspondence constraint: IDENT-IO-(F_0). Since the F_0 value in word-initial position is heavily influenced by the phonation contrast, it is likely that IDENT-IO-(FINALF_0) suffices to do the job. This constraint says that the tone value of the final temporal span should be identical for allotones in the output. Furthermore, I have demonstrated above that the optimal output form for Tone 3 (334) in initial sandhi position is a high-rising tone (45) due to the requirement of contour preservation. So I stipulate a constraint *INITIAL-34 to derive the alternation: 334 → 45.

(39) Overkill in pitch excursion

Input: 334	*INITIAL-34	IDENT-IO-(FINALF_0)
☞ a. ($_{\text{Modal}}$45, $_{\text{Murmur}}$25)		
b. ($_{\text{Modal}}$45, $_{\text{Murmur}}$23)		*!
c. ($_{\text{Modal}}$34, $_{\text{Murmur}}$24)	*!	

(where (T_1, T_2) = a pair of allotones.)

It can be seen from the above tableau that under the pressure from IDENT-IO-(FINALF_0), Tone 4 (223) is in turn forced to become a steeper rising tone, 25.

With these discussions in mind, I turn in §4.5.4 to cases in which contour preservation is apparently not obeyed in sandhi initial position, namely, Tone 1/2. In addition to this, the (non-)contouricity agreement is also not attested in Tone 1/2. I will show that a specific property of Tone 1/2 motivates these anomalies.

4.5.4 Tone 1/2 as Defective Tone

This section deals with the seemingly anomalous behaviors of Tone 1/2 (remember the "Tone 1/2 vs. Tone 3/4/5/6" distinction, *passim*). I will argue that the very distinction simply falls out in view of Tone 1/2 being *defective tones*. The asymmetry in question is repeated below.

(40) Contour flattening: Tone 1/2 vs. Contour preservation: Tone 3/4/5/6

Tone 1/2	Tone 3/4	Tone 5/6
551 → 33	334 → 45	31 → 32
231 → 22	223 → 25	221→ 22[1]

Notably, Tone 5 (31) and Tone 6 (221) have minor falls than Tone 1 (551). One might speculate that since Tone 1 (551) has the greatest pitch excursion among falling tones, this particular tone may more opt for contour leveling in a shortened rime duration under the assumption that "the greater the pitch excursion, the longer duration it requires" (Zhang 2002: 25). However, the empirical data do not seem to corroborate this assumption. First, Tone 2 (231) has roughly the same pitch excursion as Tone 5 (31), the only difference lying in the phonation contrast (cf. Figure 4-1 and Figure 4-2). Second, and more importantly, Tone 1's duration is *not* most diminished in sandhi; instead, isolation forms have the *shortest* phonetic length.

(41) Mean duration of Tone 1 in different contexts

Citation	Initial sandhi
551	33

Citation	Initial sandhi	Remarks
168.5 ms (SD=22.4)	Pattern A: 198.3 ms (SD=33.9)	citation ≈ initial
	Pattern B: 250.2 ms (SD=40.2)	***citation < initial***
	Pattern C: 259.9 ms (SD=47.3)	***citation < initial***

From (41), we can draw the conclusion that it is unlikely that a sharp fall like 551 is banned on a 250 ms-long rime. This is simply because this steep fall appears on monosyllables in isolation whose average rime duration is 168.5 ms. In addition to this, recall from the preceding section that the word-initial sandhi form of Tone 4 (223) is a low-to-high rising tone: 25. This rising tone can be licensed by a 242.3 ms-long rime ((SD=27), Pattern J; note also that 25's phonetic length is even shorter in Pattern I: 231.4 ms (SD=35.4)). It is well-known that rising tones are more marked than falling tones. This contour tone markedness hierarchy is generally translated into OT by the intrinsic ranking: *RISE » *FALL (» *LEVEL) (e.g. Yip 2002, among many others). So, if 250 ms is good for a steep rising tone, it is unexpected that a falling tone has to be completely flattened under the same environment. This is reminiscent of Zhang's (2002) 'contour reduction + rhyme lengthening' diagnosis for Hausa. This analysis is not applicable here because the case in Hausa is such that the same contour is realized on *different* types of syllable: CVV vs. CVO.

In the case of SX, syllable type is not at issue here: both Tone 1 (551) and Tone 6 (223) are long non-checked tones, i.e. tones on sonorant-final syllables. Furthermore, the prosodic position is the same: the initial position in disyllabic lexical compounds, indicating that Tone 1/6 have equivalent metrical prominence. So what we have witnessed here is the case in which *all else being equal* (as I have mentioned above), a less complex tone (high falling tone) is banned in the environment where a more complex tone (high rising tone) is instead free to occur.

Now it is time to explain why Tone 1/2 are different. The gist of my claim is that Tone 1/2 are defective tones. Phonetically speaking, Tone 1/2 are falling tones but they do not have a final target. Some remarks follow: throughout the discussion in this dissertation, I have assumed that tones are normally comprised of two "slices." Presumably, each slice is tonally specified. This kind of structure is has been labeled as "full-fledged tone" (42)a. In contrast, defective tones are tones lacking full specification (42)b. For Tone 1/2, the second slice, or the final portion of the contour, is not specified.

(42) a. Full-fledged tones b. Defective tones

With this assumption in mind, let us now go back to the thorny issue: why do only Tone 1/2 undergo complete contour reduction, while the other long tones resist doing so? In light of relational correspondence, contour preservation is essentially motivated when relational correspondence constraints are ranked over tonal markedness constraints. For full-fledged tones, the relation between temporal spans, or slices, must not be altered because of the high-ranked

RelCorr constraints. For instance, the tonal slices of Tone 3 (334) stand in the 'smaller than' relation underlyingly. If flattened out in the output under the pressure from *Rise, the now-flattened candidate incurs a violation of RelCorr$(x<y)_{\text{Nuc}}$ because level tones are characterized by the 'equal to' relation (x=y). We simply need to rank RelCorr$(x<y)_{\text{Nuc}}$ over *Rise and this partial ranking ensures that the rising contour is retained in the output. Likewise, the falling contours in Tone 5 (31) and Tone 6 (221) are also faithfully realized provided that RelCorr$(x>y)_{\text{Nuc}}$ outranks *Fall. By contrast, if there is no final specification for Tone 1/2, it follows that the RelCorr constraints are always *vacuously satisfied* because no relation is established in the input. Therefore, contour preservation is irrelevant to Tone 1/2. Under this approach, one point of importance must be addressed before we move on.

We have seen that the phonetic realization of Tone 1 in isolation is a steep fall, 551. If the pitch moves from the highest part of the F_0 range to an unspecified region (i.e. the second half), a smooth fall, e.g. 553, would be expected, instead of the actual surface form, a steep fall. It is conceivable that a more pronounced fall 551 is dispreferred: pitch movement is not supposed to be more excessive than absolutely necessary. Especially in the absence of specification, a steep fall seems even more unmotivated.

It is important to note that the contour shapes of Tone 1/2 become falling in isolation and in sandhi final position. In other words, the final fall can be interpreted as domain-final effect. That is, the final falling is due to a low boundary tone L%.

(43) Domain-final effect

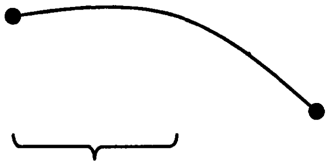
UR → SR

Tone

[Initial: 5] [Final:] L% → 551

In addition, we have seen that Tone 1/2 are flattened in *non-final* position: 551/231 → 33/22 in initial sandhi position. We can see from Figure 4-1 that Tone 1 (551) has a plateau in the first half of the contour (as for Tone 2, the initial dip is due to murmur). Therefore, it might well be the case that the first portion of the contour is truncated in non-final position, as graphically illustrated below.

(44) Initial truncation of Tone 1/2 in non-final position



Finally, the remaining issue of concern is why Tone 1 (551) and Tone 7 (55) are lowered to the midrange in initial sandhi position.

(45) No high level tone in sandhi initial position

	Citation	Sandhi initial
Tone 1	551	33 (*55)
Tone 7	55	33 (*55)

There is no apparent reason why high level tone (55 and 55) is banned in sandhi initial position. Avoidance of neutralization does not seem to play a role. In particular, for Tone 7/8, 55-23 should form a better contrast as opposed to the actual forms, 33-22, in the checked sandhi tone inventory. Due to lack of the relevant phonetic data, I leave this issue open, awaiting further research.

Before we leave this section, one important point is worthy addressing in the face of more general issues. The role of duration in contour tone licensing seems overestimated (Gordon 2001 and Zhang 2002, *inter alia*). At this point, it looks like that duration (and phonetic correlates thereof) becomes, at least in SX, orthogonal to the alleged tonal complexities as long as some threshold value on contour tone-bearing ability is satisfied.[13] For concreteness, one representative example is provided. First, it is evident that the data below clearly show the mirror image of the central tenet advocated in Gordon (1999) and Zhang (2002): contour tone is crucially licensed by the duration of the sonorous rime.

(46) The *smoother* the pitch excursion, the longer duration it requires

Tone1	Pitch value	Duration
Isolation	551	168.5ms (≈ 150ms)
Initial	33	236.1ms (≈ 250ms)

Let us simply assume that the durational difference is 150 ms vs. 250 ms. According to Zhang's (2002) system, the following constraints are intrinsically ranked: *551-ON-150ms » *551-ON-250ms. As Zhang puts it, this is because a greater pitch movement requires more time to complete it. Since a steep fall is not banned on a 150-ms long rime, we need to rank IDENT-(551) over *551-ON-150ms, as shown in the following tableau.

---

[13] Of course, one may contend that the contour-specific duration predicted by a system like Zhang (2002) is simply a *necessary* condition for contour tone licensing. For example, it may not be surprising to see that level tones are longer than falling tones, because "[p]rovided these conditions are met, languages are free to impose additional requirements on a subset of the tones in the inventory" (Zhang 2004, fn. 9). That being the case, we then want to ask: where do those "additional requirements" come from? It looks like that nothing independently predicts phonetic length in this vein. What is worse, turning back to SX, we have known that a 168.5 ms long sonorous rime can bear a steep fall. Suppose that there is language *L* whose contour tone-bearing ability is the same as SX. It wrongly predicts that the following tonal inventories may emerge: i) contour tones occur on a short vowel and level tones occur on a long vowel, or ii) contour tones are only attested on checked syllables whereas level tones occur with non-checked long syllables only. The line of reasoning is essentially the same: since a short rime is able to license a contour tone, there is nothing that can prevent level tones (rather than contour tones) from being awarded extra length. To the best of my knowledge, this hypothetical case is unattested.

(47) Steep fall on a short rime

551, 150ms	IDENT-(551)	*551-ON-150ms	*553-ON-250ms
☞ a. 551, 150ms		*	
b. 553, 150ms	*!		

However, when the rime is lengthened, the prediction is that the underlying contour should be realized at least in some fashion. In actuality, however, the contour undergoes complete contour reduction in this environment.

(48) Steep fall resists being completely flattened

551, 250ms	IDENT-(551)	*551-ON-150ms	*551-ON-250ms
!☞ a. 551, 250ms		*	*
☹ b. 33, 250ms	*!		

As mentioned earlier, the threshold value for contour tones is around 150 ms. Beyond this limit, theoretical machineries entirely based on the effect of tonal complexities on the duration seems neither necessary nor sufficient, at least in SX.

This completes our discussion on tonal realization in initial sandhi position. The next issue I would like to discuss is the tonal alternations between two adjacent tones.

4.5.5 Tonal Register (Dis)Harmony

In this section, let us look into the final issue of SX tone sandhi: tonal register (dis)harmony, a phenomenon whereby the (non-)contouricity of the preceding tones interacts with the tonal register of the following tones in an unusual way. I would like to reiterate that tonal register is not phonation register. Instead, H- and L-register here refer to higher and lower part of the F_0 range. The cutoff point of the tonal registers is around -0.5 in LZ. To begin, the generalization is repeated as follows.

(49) Conditions on Tonal Register (Dis)Harmony

a. If σ_1 is a level tone, σ_2 is H-registered (shaded cell).
b. If σ_1 is a contour tone, the register of σ_2 agrees with σ_1.

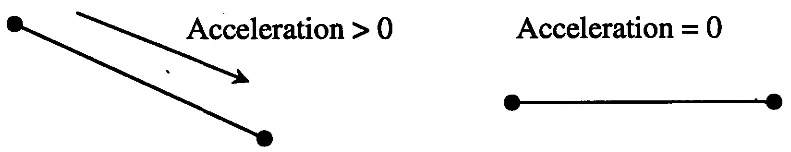
T_2 T_1	551	231	334	223	31	221	55	23
551	*33-42*	*33-42*	33-44	33-44	33-44	33-44	33-44	33-44
231	*22-42*	*22-42*	22-33	22-33	22-33	22-33	22-44	22-44
334	45-51	45-51	45-51	45-51	45-51	45-51	45-53	45-53
223	25-51	25-51	25-51	25-51	25-51	25-51	25-53	25-53
31	32-21	32-21	32-23	32-23	32-21	32-21	32-34	32-34
221	2ꜜ2-ꜜ2ꜜ1	2ꜜ2-ꜜ2ꜜ2	2ꜜ2-23	2ꜜ2-23	2ꜜ2-ꜜ2ꜜ1	2ꜜ2-ꜜ2ꜜ1	2ꜜ2-23	2ꜜ2-23
55	*33-42*	*33-42*	33-33	33-33	33-33	33-33	33-44	33-44
23	*22-42*	*22-42*	22-33	22-33	22-33	22-33	22-44	22-44

Several things are worth noticing from the table in (49). First, the lowest F_0 values of tones preceded by a level tone (marked with a shaded cell) are invariably above 3 (at the five-point scale, or above -0.5 in LZ). In other words, these final sandhi tones all elevate to H-register. This is primarily due to the fact that murmur disappears in non-initial position so that the once-depressed F_0 curves "bounce back" to H-register, assuming that H-register is the default register for modal phonation (cf. Yip 1993). That being the case, we do not expect L-register tones on the final syllables. Second, if the preceding tone is a contour tone, it appears that the right edge of the preceding tone and the left edge of the following tone must be agree in tone level. In particular, if the initial tone has a L-register tonal offset (e.g. Tone 5 (31) and Tone 6 (221)), it is observed that the following tone must start at the same tone level tone as well. Finally, the present phenomenon raises at least two serious challenges to the following standard analyses. On the one hand, the ranking IDENT-INITIALTONE » *L-REGISTER » *H-REGISTER » *TONE predicts that H-register is the default register and furthermore, L-register is banned on non-initial syllables. As it stands, this accounts for the generalization in (49)a but *not* in (49)b: L-register is attested on non-initial syllables. On the other hand, we can formulate an AGREE[F] constraint (or Hyman and VanBik's (2004) NoJump) to penalize an abrupt F_0 transition across syllable boundary, i.e. the right edge of the preceding tone and the left edge of the following tone have the same pitch value. But this time, it turns out that the generalization in (49)b but *not* in (49)a is predicted. I.e. an abrupt pitch jump is not impossible across syllable boundary. In sum, it seems unlikely that tonal register (dis)harmony is not motivated by some all-purpose tonal markedness consideration, be it default tonal register or minimization of F_0 fluctuations across syllable boundary.

The key to the solution lies in the observation that it is contour tones, but not level tones that are able to change the tonal onset of the following syllable. This is what Xu (1997) termed "exclusive carry-over assimilation." In Xu's (1997) model, pitch movement is decelerated when a target is being approached. In light of this, the phenomenon in question can be attributed to minimization of articulatory effort. The idea is that pitch movement of level tones is not accelerated because the pitch targets are at the same tone level. In contrast, the tonal contour dynamically moves from the initial target to the final target in a contour tone. In other words, pitch movement must be accelerated, upwards or downwards, to achieve the final target. The point is illustrated in the schematic daigrams below.

(50) Pitch movement

Contour tone: Acceleration > 0 Level tone: Acceleration = 0



I postulate that if the minimization of articulatory effort constraint is active, the tension of the responsible laryngeal muscles will not be relaxed,[14] when the final target is being approached. One of the desirable results is that when the initial target of the following tone is an unlike one, the glottis is prevented from shifting from one state to another. Therefore, laryngeal activities are not adjusted to reach the next target that is distinct. The discussion can be formalized by the following constraint:

(51) MINIMIZE EFFORT

> "The F_0 value of the offset of *contour* tone T_1 and the onset of T_2 should be within the same range."

Notice that by "the same range," I mean a specific threshold LZ value for level tones (cf. (22)). The following tableau shows how this superficial spillover effect works.[15]

(52) 31-31 → 32-21

31-31	MINIMIZE EFFORT (*32-3)	IDENT-FINALT
☞ a. 32-21		*
b. 32-31	*!	

As we can see, in the winning candidate (a), the tonal offset of the preceding syllable extends into the tonal onset of the following syllable, resulting in a smooth F_0 transition. This tone encroachment is motivated due to the inertness of laryngeal activities during the transition from one target to another. As a result, an abrupt F_0 transition is penalized by the top-ranked MINIMIZE

---

[14] For F_0-related laryngeal muscles activities, see Hirose *et al.* (1970), Ohala and Hirose (1970), Garding *et* al. (1970), Hirose and Gay (1972), Collier (1975), Erickson (1976), Atkinson 1978, Harris (1978) for the cricothyroid muscle and (Erickson 1976) and Atkinson (1978) for the vocalis (VOC) muscle. See also Hirose (1997) for a useful overview.

[15] Notice that the motivation here is very different from the conventional spillover effect (e.g. Ohala 1978). This is because the canonical duration of contour tones, as extensively discussed above, does not play a significant role in the present case.

EFFORT constraint, at the cost of unfaithfully realizing underlying tone specification. Candidate (b) is ruled out because laryngeal activities are adjusted to reach the onset target of T_2.

On the other hand, consider now the following case in which the second tone raises to H-register (cf. (49)a). We have known that level tones are not specified with the ACCELERATION feature, so it is predicted that the second syllable must be invariably H-registered if *L-REGISTER is ranked over *H-REGISTER.

(53) "Default" tonal register: 23-551 → 22-42

23-551	MINIMIZE EFFORT	*L-REGISTER	*H-REGISTER
☞ a. 22-42			*
b. 22-22		*!	

In conclusion, the present data indicate that we should also incorporate articulatory constraints in the analysis of tone sandhi. As shown, the tonal register (dis)harmony phenomenon cannot be adequately handled in terms of the known approaches.

4.6 Conclusion

I summarize the main points made in this chapter. We have presented arguments establishing that the non-equal to relation is not reducible to the other relations in the relational correspondence theory, albeit counterintuitive as it appears to be. The evidence comes from the phenomenon I termed (non-)contouricity agreement. Based on the fine-grained phonetic description, I have provided a formal analysis of the disyllabic tone sandhi of SX, appealing to relational correspondence. I have also demonstrated that articulatory and perceptual principles must factor into a satisfactory account for a complicated tone sandhi system like SX.

Chapter 5 Slope Correspondence

5.1 Licensing by Slope

In the tone literature, it has long been noted that contour tone licensing crucially hinges on the first two dimensions in (1).

(1) Dimensions of Contour tone licensing

i. Syllable quantity (mora count) or duration of sonorous rime
ii. Neighboring tones (e.g. Hyman's (to appear) "the law of the like neighbor")
iii. Slope (The goal of this chapter)

The core issue of this chapter is to investigate an understudied dimension of contour tone licensing: slope. Slope refers to the degree of the steepness of a straight line. As a working hypothesis, a tone can be described through an idealization according to which pitch contours are regarded as a straight interpolation between the F_0 maximum and the F_0 minimum. Slope describes the ratio: F_0 difference over duration. Moreover, in a famous summary of early work on tone perception, Gandour (1978), citing Gandour and Harshman (1978), has already pointed out that slope serves as one of the five dimensions for tonal or atonal language speakers to distinguish one contour tone from another. Nevertheless, there is little attention to the utilization of 'slope-matching' as a theoretical diagnostic of tone sandhi.[1] To fill this gap, I consider tonal faithfulness in terms of slope in this chapter.

---

[1] As a precursor of slope correspondence, Zhang's (2002b) formulation of the PRESERVE(tone) constraint family is an exception. As we will see, the size of the F_0 difference is essential in the measurement of the steepness of a tone. At first sight, slope correspondence may not be empirically distinguishable from PRESERVE(tone). One important distinction, however, lies in that slope correspondence is not restricted to the syllable domain, while PRESERVE(tone) only assesses tonal identity within syllables. See also §5.5.

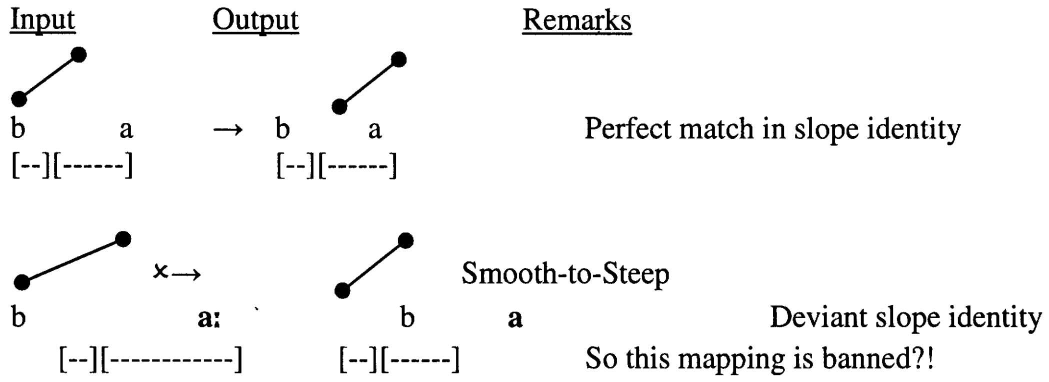
With regard to the precursor of research projects along this line, we have learned in §1.2 that Steriade (2006) termed the dimension of inquiry "slope correspondence" and proposed the constraint MATCH-SLOPE to derive a variety of effects in sonority. The formulation of Steriade's (2006) MATCH-SLOPE is repeated as follows.

(2) Steriade's (2006) MATCH-SLOPE

D is an auditory dimension.
Let *xy* be a sequence of elements in S_1, where *x* precedes *y*
Let *ab* be a sequence of elements in S_2, where *a* precedes *b*
x is the S_1 correspondent of *a*;
D(*x*)=the D value of *x*; D(*a*)=the D value of *a*; etc.
Difference between $\Delta(D(x)-D(y))$ and $\Delta(D(a)-D(b))$ does not exceed n.

Remarkably, the duration of two successive elements is not taken into consideration in Steriade's MATCH-SLOPE. It is fair to say that there would be no harmful repercussion when slope correspondence is invoked in an auditory dimension such as sonority contour and metrical prominence. To see why, let us consider sonority contour for example. Suppose that there are two stop-vowel sequences: *ba* and *baː*. If we take the midpoint of a vowel as the "element" in (2), it appears that there is a slope difference between *ba* (steeper) and *baː* (shallower), as graphically represented below.

(3) Unattested mapping: Vowel shortening is blocked by slope mismatch



Suppose further that long vowels may undergo shortening on the surface. It should be reasonable to say that slope mismatch between the sonority contours in the Input-Output relation is one of the least plausible motivations for why a long vowel is prohibited from being shortened in the output. I.e. we do not find vowel shortening blocked to maintain slope. What really matters in sonority contour seems to be the *syntagmatic* difference (on a perceptual scale) between two immediately adjacent elements, rather than the phonetic length between them. By contrast, F_0 is perhaps the only auditory dimension in which duration is a vital factor in slope correspondence. As mentioned earlier, rime duration (or mora count) has long been noted to play an important role in contour tone licensing. Intuitively, we can imagine that if an underlying sharp rise is to be realized on a durationally shorter rime, this sharp rise may be reduced to a level tone to avoid a big pitch excursion on a short time span for the sake of minimization of articulatory effort. So slope identity may be incomplete in the output: the perceptual difference between a sharp rise and a level tone is supposed to be drastic. To maintain slope identity, pitch excursion may be enlarged or *rime duration* can be lengthened (see §5.2.1 for more discussion). It then appears that pitch excursion alone is not sufficient in characterizing slope difference. Steriade's MATCH-SLOPE must be modified before it can be incorporated into the model of slope correspondence in tone sandhi because the MATCH-SLOPE constraint in (2) does not provide a good basis for comparing the difference of tones of different slopes.

5.2 Introducing Slope Correspondence

5.2.1 Basics of Slope

Slope describes the steepness of a straight line. A higher slope value indicates a steeper incline. Slope is defined as the *ratio* of the altitude change to the horizontal distance between any two points on the line.

The slope of a line in the plane containing the x and y axes is conventionally represented by the letter ***m***, and is defined as the change in the y coordinate divided by the corresponding change in the x coordinate, between two distinct points on the line. This is described by the following equation:

$$(4)\ m = \frac{\Delta_y}{\Delta_x}$$

Given two points (x_1, y_1) and (x_2, y_2), the change in x from one to the other is $x_2 - x_1$, while the change in y is $y_2 - y_1$. Substituting both quantities into the above equation obtains the following:

$$(5)\ m = \frac{y_2 - y_1}{x_2 - x_1}$$

Turning back to tone, we can say that the change in x pertains to the rime duration, while the change in y means the difference from the starting point of F_0 to the endpoint of F_0. Assuming that tone T has two pitch targets (t_j and t_i, where t_j precedes t_i) and the interpolation between the two pitch targets is idealized as a straight line, the slope of tone T (call it m_T) is defined as the following ratio.

$$(6)\ m_T \quad = \quad \frac{F_0(t_i) - F_0(t_j)}{\mathrm{Time}(t_i) - \mathrm{Time}(t_j)}$$

For a better understanding, consider now the following two rising tones. Let us suppose that the rime durations of the two rising tones are identical and that their starting F_0 values of them are also the same, namely, 100 Hz. The only difference lies in the value of the F_0 endpoint. Tone 1 is 100 Hz higher than Tone 2 in the final pitch value (= 300 Hz – 200 Hz). It should be obvious that Tone 1 is steeper in slope than Tone 2 since the change in the overall F_0 contour is greater in Tone 1. The slope of Tone 1 (m_1) and Tone 2 (m_2) can be obtained according to the equation in (6). Since 2 is greater than 1, Tone 1 is steeper than Tone 2.

(7) Slope of rising tone

	Beginning point	Endpoint	Rime duration
Tone 1	100 Hz (t_j)	300 Hz (t_i)	100 ms
Tone 2	100 Hz (t_q)	200 Hz (t_p)	100 ms

$$m_1 = \frac{F_0(t_i) - F_0(t_j)}{\mathrm{Time}(t_i) - \mathrm{Time}(t_j)} = \frac{300\,\mathrm{Hz} - 100\,\mathrm{Hz}}{100\,\mathrm{ms}} = 2$$

$$m_2 = \frac{F_0(t_p) - F_0(t_q)}{\mathrm{Time}(t_p) - \mathrm{Time}(t_q)} = \frac{300\,\mathrm{Hz} - 200\,\mathrm{Hz}}{100\,\mathrm{ms}} = 1$$

Let us now turn to falling tone. In (8), the starting F_0 and ending F_0 values given in (7) are swapped while the rime duration remains intact. Intuitively, Tone 3 is steeper in slope than Tone

4 since the F_0 change is greater in Tone 3. According to the equation in (6), the slope of Tone 3 is -2 and the slope of Tone 4 is -1. Here we take the falling tones' absolute value to ensure that the slope of Tone 3 is greater than that of Tone 4: $|-2| > |-1|$. In other words, the larger the absolute value of a slope is, the steeper a contour tone is.

(8) Slope of Falling Tone

	Beginning point		Endpoint		Rime duration
Tone 3	300 Hz	(t_j)	100 Hz	(t_i)	100 ms
Tone 4	200 Hz	(t_q)	100 Hz	(t_p)	100 ms

$$m_3 = \frac{F_0(t_i) - F_0(t_j)}{\text{Time}(t_i) - \text{Time}(t_j)} = \frac{100\,\text{Hz} - 300\,\text{Hz}}{100\,\text{ms}} = -2$$

$$m_4 = \frac{F_0(t_p) - F_0(t_q)}{\text{Time}(t_p) - \text{Time}(t_q)} = \frac{100\,\text{Hz} - 200\,\text{Hz}}{100\,\text{ms}} = -1$$

Finally, I assume that the slope of level tones is approximately zero (0). This is of course a simplification but in reality one hardly sees a level tone without any F_0 fluctuation (e.g. 100 Hz throughout the entire contour).

This completes our introduction to the measurement of the slope for tone. In the following section, I discuss the metric used for comparisons of slope similarity throughout this dissertation: the Index of Slope Difference.

5.2.2 Index of Slope Difference

Provided that the slope correspondence constraint assesses identity of correspondent slopes, a technical issue arises. Given two pairs of correspondent slopes, $m_A \rightarrow m_B$ and $m_A \rightarrow m_C$, how do we know which slope mapping is more faithful (or similar)? Recall that slope is a ratio. So an adequate metric to evaluate the degree of deviation is to obtain the quotient of two slopes in comparison. This metric is what I termed Index of Slope Difference (abbr. ISD). Supposing that slope m_B is the output/derivative correspondent of slope m_A, the ISD is calculated according to the following equation.

(9) Index of Slope Difference (ISD) = $$\log\left(\frac{\text{Input/Base Slope}}{\text{Output/Derivative Slope}}\right) = \log\left(\frac{m_A}{m_B}\right)$$

The reason why I take logarithm (to base 10), rather than the ratio on its own, for the ISD is simply for ease of comparison. To see why, I now illustrate a simple application of the ISD. As we can see below, the ISD for a "perfect match" in slope is zero (0). If a slope mapping is steep-to-smooth, the ISD is greater than zero (0). If a slope mapping is smooth-to-steep, the ISD is smaller than zero (0), i.e. of the negative value. This means that the 'reference point' of this similarity index shifts to zero (0), rather than one (1). In other words, after the logarithmic manipulation, slope similarity is manifested in a more straightforward fashion.

(10) Index of Slope Difference

	Input → Output	ISD
Perfect match in slope (ISD=0)	m_A (=4) → m_B (=4)	log (4/4) = log 1 = 0
Steep-to-Smooth (ISD>0)	m_A (=4) → m_C (=2)	log (4/2) = log 2 = 0.3
Smooth-to-Steep (ISD<0)	m_A (=4) → m_D (=5)	log (4/5) = log 0.8 = -0.09

Turning back to the question raised at the outset, how do we know which mapping of correspondent slopes, e.g. $m_A \rightarrow m_B$ or $m_A \rightarrow m_C$, is more similar? From (10), we draw the conclusion that the mapping whose ISD is closet to zero (0) is most faithful in slope, namely, $m_A \rightarrow m_B$. In this mapping, the underlying slope is 4 and the surface slope is also 4. This means that slope is replicated in the output (but notice that duration or pitch excursion may vary drastically).

The ISD value of this perfect match is zero (0) according to the above equation. So if two pairs of correspondent slopes are being compared, the ISD is a reliable means of degree of similarity.

More technical issues must be addressed before we leave this section. First, the ISD cannot be used for comparisons of slopes with reversed direction. This is because the slope for falling tones is negative whereas rising tones have a positive slope. By definition, the number for a logarithm *cannot* be negative (i.e. the quotient of two slopes in comparison). This would not be a serious problem since contour direction is an essential dimension for the perception of contour tones (Gandour 1978). That being the case, it should be reasonable to assume that slope identity does not (perhaps cannot) play a role in mappings of tones with reversed contour direction. Under this view, it is not possible to tell whether a mapping from a steep fall to a steep rise is more similar than a mapping from a steep fall to a smooth rise. I assume that this is the job of the Input-Output faithfulness.

Second, since log 0 is by definition negative infinity, level tones cannot be zero (0) in slope. Otherwise, we are not able to obtain an ISD value if a level tone maps to a contour tone (or vice versa). Again, this should be a harmless stipulation because level tones have, if not always, some tiny fluctuations on surface.

This completes our introduction to the Index of Slope Difference. In the following section, I set out to formalize the slope correspondence constraint for tone sandhi and discuss the essential assumptions made in this theoretical apparatus.

5.2.3 Formulating the Slope-matching Mechanism

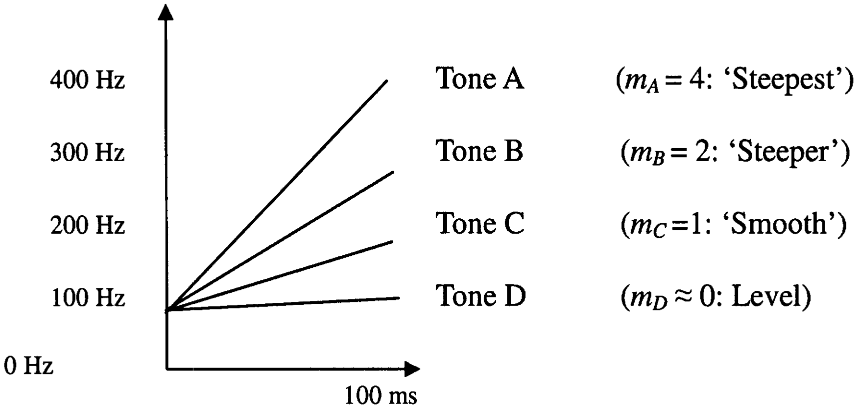
The central tenet of relational correspondence is that phonological processes are shaped by pressure to maintain perceptual similarity between correspondent relations of successive elements. As we have discussed, slope is defined as the ratio between the altitude change (i.e. F_0

difference) and the horizontal difference (i.e. rime duration). Importantly, I assume that F_0 difference and duration are obtained between the two slices contained in Tone T. In other words, slope correspondence is a specific instantiation of relational correspondence. So in addition to the four contour correspondence relations discussed in chapters 2 to 4, slope is subsumed as an independent relation in the relational correspondence constraint family.

5.2.3.1 Defining Match-Slope

The formalization of slope correspondence is couched in Steriade's (2001) P-map approach. Confusability is inversely proportional to the degree of distinctiveness of contrasts: the greater the degree of confusability, the more similar two elements are. In the same vein, I assume the premise such that slope-matching assesses the relative degree of similarity. We have learned that, at least for tone, the source of similarity must come from linguistic knowledge that is tuned up on a language-specific basis. For example, a 10 Hz rise or fall in pitch may or may not be linguistically relevant for language *L*. Furthermore, a 10 Hz interval in the higher or the lower part of the pitch range may also have different bearings on tonal distinctiveness (i.e. the positive skew of F_0 distribution in Chapter 4). Aside from these language-specific similarity factors, however, slope-matching seems to behave in a consistent manner. The first evidence comes from 'faithfulness to the marked' in tone sandhi. Let us first understand the phenomenon in the abstract. A schematic illustration is provided below.

(11) Hypothetical slope categories (cf. (4))



Intuitively, the most faithful correspondent slope for Tone A should be Tone B. This is rightly so because the slopes of Tones C and D deviate more from that of Tone A. It is reasonable to conjecture that Tone B is 'one step away' from Tone A, Tone C is 'two steps away' from Tone A and Tone D is 'three steps away' from Tone A on a hypothetical perceptual scale. Suppose that there are two rising tones in an inventory, Tone **A** and Tone **C**. Interestingly enough, it is attested in Hangzhou Chinese that a steep rise Tone A maps to a smoother rise Tone B in sandhi and Tone C is completely flattened, becoming Tone D in the same context. The phonetic data below are based on my acoustic measurements (see §5.3 for more details). As we can see, the sharp rise 13 is faithfully rendered in sandhi. The slope difference is tiny. In contrast, the smooth rise is completely reduced to a level tone. In terms of (11), Tone A=13, Tone B=13', Tone C=23 and Tone D=22.

(12) Preservation of the marked in Hangzhou Chinese

Tone				$\Delta F_0$10%-90%	Rime duration	Slope
L-register	Citation	13	(=A)	65.1 Hz	341.2 ms	0.19
Sharp rise	Initial	13'	(=B)	49.5 Hz	282.5 ms	0.17
L-register	Citation	23	(=C)	36.9 Hz	425.9 ms	0.08
Smooth rise	Initial	**22**	(=D)	4.6 Hz	323.4 ms	0.01

More importantly, a non-trivial issue is the following: why is the pitch excursion size (i.e. ΔF_0) *not* faithfully realized in sandhi? In particular, sandhi tone 13' has a phonetic length of 282.5 ms. This is more than enough time for a 65.1 Hz pitch rise because the minimum time to implement a 65.1 Hz rise is 138.6 ms (=100.4+5.8*6.6st), according to Xu and Sun (2002). Therefore, there is no apparent reason why IDENT-ΔF_0 is not active. This constraint requires that pitch excursion size should be identical in the input and the output. In view of slope, it may not be difficult to understand why this could be the case. This is because if IDENT-ΔF_0 were active, the sandhi slope would be steeper: 65.1 Hz / 282.5 = 0.23. Recall that the slope of tone 13 is 0.19. This simple calculation reveals two important properties of slope-matching. First, *slope matching is more important than faithful realization of the ΔF_0 (or pitch target-matching).* This observation can be captured by the ranking below.

(13) MATCH-SLOPE » IDENT-ΔF_0

Second, smooth-to-steep slope matching seems to be dispreferred. It can be seen from (12) that sandhi slopes are shallower than citation slopes. The question is why a steeper slope does not fare better in sandhi? Importantly, since duration is diminished in sandhi, a steeper slope means that IDENT-ΔF_0 (i.e. pitch target-matching) is more satisfied. In spite of this, the actual sandhi slope is shallower.

Taken together, the above discussion indicates that the following mappings are rarely found, if not absent at all.

(14) Seemingly unattested base-derivative mappings of two rising tones (steep vs. smooth)

iv. The steep rise becomes a level tone and the smooth rise is preserved.
v. The steep rise becomes a yet steeper rise and the smooth rise becomes a steeper rise.

Some discussion is in order. Duration is generally compressed in connected speech. So given the well-established fact according to which it takes more time to implement a contour tone with greater pitch excursion (Gordon 2001, Zhang 2002, *inter alia*), it is expected that a sharp rise, rather than a smooth rise, would be more 'vulnerable' in context. Therefore, as stated in (i) above, the more marked steep rise is supposed to undergo complete contour reduction, while the contour of the smooth rise should be retained. However, this prediction does not gain support from the empirical facts. In actuality, it is the steep rise that resists contour leveling and the smooth rise is flattened out, at least in Hangzhou Chinese. It should be obvious that employing the conventional markedness constraint is incapable of dealing with the 'markedness reversal' problem (cf. de Lacy 2002a). We need the following ranking of constraints *SMOOTHRISE » IDENT-(TONE) » *STEEPRISE to derive the desired result for which the steep rise is faithfully realized, whereas the shallow rise becomes a level tone. In spite of this seemingly stipulative analysis, I will argue that a more promising solution resides in faithfulness, or more specifically, slope correspondence. To see how this works, the MATCH-SLOPE constraint is firstly formulated as follows (cf. Boersma's (1998) *REPLACE, Zhang's (2002) PRESERVE-(tone), or Zuraw's (2005) *MAP).

(15) MATCH-SLOPE$_{\text{ISD}\rightarrow 0}$

Tone *i* has two correspondents Tone *j* and Tone *k*.
Let the slopes of Tone *i/j/k* be $m_{i/j/k}$.

The Index of Slope Difference (ISD) $= \log\left(\frac{m_i}{m_j}\right) = p$; $\log\left(\frac{m_i}{m_k}\right) = q$.

Don't map Tone *i* to Tone *k*, if $q > p > 0$.

Suppose that there are three tones, Tones *i*, *j*, and *k* and Tone *i* > Tone *j* > Tone *k* in steepness. Active MATCH-SLOPE$_{\text{ISD}\rightarrow 0}$ predicts that Tone *i* preferentially maps to Tone *j*, rather than Tone *k*. More precisely, in the preceding section, the Index of Slope Difference (ISD) was proposed as a similarity index of correspondent slopes. Since Tone *i* and Tone *j* are more similar in slope, the ISD of the mapping from Tone *i* to Tone *j*, *p* is closer to zero (0) than that of the mapping from Tone *i* to Tone *k*, *q*. Recall that a zero (0) ISD means a perfect match in slope. Consequently, if slope identity is more stringently enforced, it follows that the 'markedness reversal' phenomenon takes place: a more 'marked' tone (i.e. of the steepest slope: Tone *i*) does not correspond to the least marked tone (i.e. of the smoothest slope: Tone *k*). To fit into the discussion in (14)i, we can imagine that Tones *i* and *j* are the steep/smooth rising tones and Tone *k* is the level tone. It then follows that MATCH-SLOPE successfully predicts the expected mapping: the steep rising tone does not map to the level tone.[2] This behavior is exactly what we have seen in the case of Hangzhou Chinese (12).

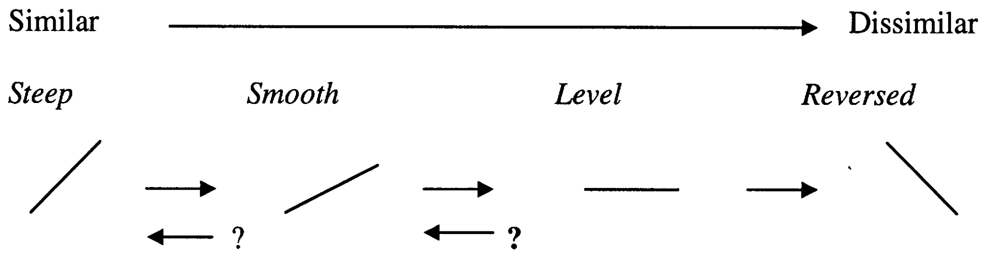
To visualize the present generalization, consider now the graphic presentation below (cf. the introductory discussion of the 'non-equal to' relation in chapter five). As it stands, provided that slope similarity is respected, the preferential retention of the contour of a steep rise is warranted:

---

[2] According to the formulation of MATCH-SLOPE$_{\text{ISD}\rightarrow 0}$, one may raise the issue of computational complexity. But notice that tone production and perception are in essence categorical. The model advanced here is able to predict an approximate pitch shift size only.

if a smooth rise output form is available, the completely flattened candidate (=level tone) loses out.

(16) Degree of similarity of correspondent slopes for a steep rising tone



Now let us turn to another seemingly unattested pattern described in (14)ii, namely, the rarity of the *smooth-to-steep* slope mappings. This statement manifests another essential property of slope correspondence, which may be termed 'the antisymmetry of slope mapping.' To my knowledge, *smooth-to-steep* slope mappings are surprisingly rare.[3] This seems to imply that steep-to-smooth slope mappings are more *biased* in slope comparisons. This observation at first blush sounds banal. But I would like to contend that it is not the case. Some discussion is in order. When the durational difference between citation and sandhi forms is large enough, it may not be possible to realize the same number of tonal contrasts in sandhi (but see fn. 4). Cross-linguistically speaking, it is more preferable to reduce the size of pitch excursion in a diminished duration, normally resulting in tonal neutralization. One extreme case is documented in Chen (2000): In New Chongming, a dialect of Northern Wu Chinese, seven (7) citation tones are distinguished in isolation but only one (1) tone appears in unstressed syllables of an accentual phrase. In contrast,

---

[3] As far as I know, there are two apparent counterexamples. But I think both of them are otherwise motivated. First, we have discussed in chapter five that Tone 4 (223) in Shaoxing Chinese changes to a steeper rising tone 25 in initial sandhi position. This is attributed to the allotone explanation. Second, the low-falling tone 21 in Taiwanese and Xiamen Chinese becomes the high-falling tone 51 in non-initial position. As is well-known, tone sandhi in these two languages (and the other varieties in the same language group) is of the *paradigmatic substitution* type (Schuh 1978, among many others). Assuming that slope correspondence is basically found only in syntagmatic contexts, it should be safe to conclude that 21 → 51 does not count as a real counterexample.

expansion of pitch excursion seems unattested in the sandhi tone inventory. That is to say, a less pronounced rise/fall in citation changes to a more pronounced rise/fall in durationally shorter sandhi syllables. The lack of 'contextual enhancement' in sandhi has been arguably attributed to articulatory difficulties (i.e. the duration-based account in (1)i). However, this generalization needs qualification. By way of a concrete example, let us consider the behaviors of the rising tone in Lhasa Tibetan. Of interest is the following asymmetry. In disyllabic lexical compounds, the LH in short syllables (CV: call it short LH) is reduced to a low level tone L in initial sandhi position. As for the LH in long syllables (CVV/CVN: call it long LH), it is also completely flattened in the same context. The gross generalization is that no contour tone is allowed in initial sandhi position. An all-purpose positional markedness constraint such as Yip's (2002) LICENCE-CONTOUR 'no non-final contour tones' is sufficient to derive the desirable results. In terms of phonetically driven tonal markedness (e.g. Gordon 2001, Zhang 2002b), however, the prohibition on non-final contour tones in Lhasa Tibetan is not expected. Consider now the relevant phonetic data below (based on my own acoustic study; see §5.4 for more details).

(17) Lhasa Tibetan: Evidence for all-purpose positional markedness?

Isolation	Initial sandhi position		
LH in short syllables	Rise: ΔF_0=30 Hz, 160 ms	→	Level: ΔF_0 <10 Hz, 80 ms
LH in long syllables	Rise: ΔF_0=40 Hz, 350 ms	→	Level: ΔF_0 < 10 Hz, 220 ms

For short LH, realizing a 30 Hz pitch rise on 80 ms is articulatorily impossible. Xu and Sun's (2002) linear equation for minimum time of pitch rise states that it takes more than 100 ms to implement any rising tone. The real challenge lies in the realization of word-initial long LH. Why isn't long LH faithfully rendered in initial sandhi position? As we will see in §5.4, a pitch excursion of 30 Hz is the *lower limit* of pitch shift for LH in this language. So, according to Xu and Sun (2002), 220 ms is more than enough time for a 30 Hz rise (where minimum time = 119.8 ms = 100.4+5.8*3.36). Moreover, in isolation short LH, 160 ms is able to accommodate a rising tone. There is then no *a priori* reason why a long syllable whose average rime duration is 220 ms, be it CVV or CVN, cannot bear a rising tone. The asymmetry receives a more straightforward explanation if slope-matching is taken into consideration. Let us suppose that a 30 Hz rise in pitch is realized on word-initial short and long syllables. Given these assumptions, the slope values for these mappings are provided as follows.

(18) Lhasa Tibetan: No smooth-to-steep slope mappings

	Isolation	Slope	Hypothetical initial rise	Slope
Short LH	ΔF_0=30 Hz, 160 ms	0.18	ΔF_0=30 Hz, 80 ms	0.37
Long LH	ΔF_0=40 Hz, 350 ms	0.11	ΔF_0=30 Hz, 220 ms	0.13

It can be seen from the above illustration that the slope values are *increasing* from isolation to sandhi, which means that a smooth rise maps to a steep rise. Mappings of this sort are disallowed because they contradict the proposed bias: a smooth slope in input/citation does not map to a steeper slope in output/sandhi, even though this mapping is articulatorily feasible. For the present case, a smooth slope could change to a steeper slope without noticeable difficulties. Moreover, recall our discussion on IDENT-ΔF_0 in (13). The hypothetical sandhi form of long LH is *more faithful* to the pitch excursion size of the isolation/input form if compared to the actual surface form, a low level tone (cf. (17)), suggesting that the dispreference for slope mappings of this sort should be otherwise motivated.

In sum, all else being equal, the slope of the inferred input (i.e. the citation tone) is preferentially *decreasing* in output/sandhi. This bias is formalized as the following slope correspondence constraint.

(19) MATCH-SLOPE$_{ISD>0}$

Tone *j* is the correspondent of Tone *i*
Let the slopes of Tone *i/j* be $m_{i/j}$.

The Index of Slope Difference = $\log\left(\frac{m_i}{m_j}\right) = p.$

Don't map Tone *i* to Tone *j*, if $p<0$.

Here I provide a plausible motivation for the constraint against a smooth-to-steep slope mapping. Given the assumption that the same syllable has a shorter duration in polysyllabic words than in monosyllabic words, realizing a steeper slope in sandhi contexts will cost more articulatory effort, especially in trisyllabic or quadrisyllabic words. Since slope identity is expected to be preserved across the board, it appears that a smoother slope fares better.

In summary, the essential characteristics of slope correspondence are recapitulated below.

- To maintain slope identity, tonal markedness may be voided. I.e. the 'faithfulness to the marked' phenomenon is better treated in terms of slope correspondence.
- Slope in the (inferred) input preferentially decreases in the output: mappings from steep-to-smooth slopes are more favorable than mappings from smooth-to-steep slopes.

The first property is formalized as the MATCH-SLOPE$_{ISD\rightarrow 0}$ constraint in (15) and the second as the MATCH-SLOPE$_{ISD>0}$ in (19). I further assume these two slope correspondence constraints are intrinsically ranked:

(20) MATCH-SLOPE$_{ISD>0}$ » MATCH-SLOPE$_{ISD\rightarrow 0}$.

In other words, smooth-to-steep slope mappings (i.e. ISD<0) are worse than steep-to-smooth slope mappings even though their ISD values are equidistant to 0. For example, -1 and 1 are equidistant to 0 (= perfect match in slope). So MATCH-SLOPE$_{\text{ISD}\rightarrow 0}$ is unable to decide the optimal output. According to the ranking above, the mapping whose ISD is 1 is selected as the winner. As we have discussed earlier, this is because correspondent slopes from steep to smooth are dispreferred.

So far I have laid out the slope correspondence constraints. Some issues, however, must be clarified before we move on to the analyses, to which I turn in the subsequent sections.

5.2.3.2 Inferred Input

It has been made explicit in the foregoing sections that slope correspondence relies on perceptual comparisons. In principle, it is actual surface forms, rather than an underlying form and a surface form that are being compared. This issue is especially important for slope-matching because the measurement of the steepness must be based on the exact duration. I would like to adopt the hypothesis of *inferred input* (Jun 2002, Steriade 1997) and I leave the issue open as to whether duration should be specified in the underlying representation. Steriade defines the inferred input as a hypothesized phonetic interpretation of the input. The input and the inferred input are not distinguishable except that all phonetic details are specified in the latter. Building on McCarthy's (1999) formalism of Sympathy Theory, Jun (2002: 13) proposes that "the inferred input be the most harmonic among candidates which obey ALL context-free IO faithfulness" (small capitals in original). Given this, slope identity to the inferred input is now accessible because phonetic length is both specified in the inferred input and a surface form. Therefore, perceptual comparisons of correspondent slopes are made possible according to the inferred input hypothesis.

In case of the Sino-Tibetan languages, I assume that the inferred input is the monosyllabic form in isolation (a.k.a. 'citation tone' or 'basic tone').[4] On the other hand, for African and other languages, it is conceivable that larger constituents such as prosodic words and/or phrases may serve as the inferred input.

5.2.3.3 Slope in Different Speech Rates

Unlike vowel reduction, tonal neutralization could be "undone" at slower speech rate (cf. Flemming 2001). One of the well-known examples comes from the allotonic distribution of Tone 2 in Beijing Chinese.[5] In isolation, Tone 2 is a high rising tone, conventionally transcribed as 35 on the five-point scale. At a relatively fast speed, however, Tone 2 becomes 55 if preceded by Tone 1, the high level tone 55 and followed by any lexical tone. Chao (1968: 27-8) termed this phenomenon 'a tone sandhi of minor importance,' probably due to the fact that this process occurs in a considerably restricted context.

---

[4] Ting (1982) observes that diachronic underlying tonal contrasts may be preserved in sandhi tones and neutralized in citation in a handful of Sinitic languages. So Ting (and other researchers) takes this as a challenge to the prevalent view according to which citation tones are the underlying tones (see Chan 1993 for more discussion). This may not be that problematic as it has been believed. Consider first the following diagram.

(i) Input
↙ ↘
Citation Sandhi
(Base) (Derivative)

Since citation/isolation and sandhi tones occur in mutually exclusive environments, they may be subject to different markedness constraints. Distinctiveness constraints determine whether a specific tonal contrast is sufficient in a given context (Flemming 2002). If not, neutralization takes place. Suppose that Tone *i* and Tone *j* are contrastive in sandhi but are neutralized to (or corresponds to) Tone *k* in isolation. Then the question is: are there two inferred inputs (i.e. Tones *i* and *j*) or just one (i.e. Tone *k*)? I do not discuss cases like this in this dissertation. But I would like to point out that it might still well be the case that Tone *k* is the inferred output. The reason is as follows. The effect of slope correspondence may be overridden between Tone *k* in citation and Tone *i/j* in sandhi, presumably due to some high-ranked tonal distinctiveness constraint for the sandhi tone inventory. Conversely, if slope correspondence is top-ranked, Tone *k* and Tone *i/j* must be similar in slope. Of course, this conjecture is subject to justification but is nevertheless reasonable, as far as I can tell.

[5] See also Ao (1993) for discussion on the interactions between speech rates and tone sandhi in Nantong Chinese, a dialect of Northern Wu Chinese.

(21) Two types of tonal neutralization in Beijing Chinese

a. 'Undoable' tonal neutralization
55-35-T → 55-**55**-T *only in fast speech rate* (T = any lexical tone)

b. Absolute neutralization (Tone 3 Sandhi)
213-213 → **35**-213

On the other hand, the famous Tone 3 sandhi in Beijing Chinese is also tonal neutralization in its own right. But Tone 3 sandhi cannot be undone even at an extremely slow speech rate. The asymmetrical behaviors seem to suggest that tonal distinctiveness may be overridden by pressure from extreme articulatory difficulties at fast speech rate (e.g. (21)a). More importantly, since slope and duration interact in a non-trivial way, does the above asymmetry indicate that slope-matching should be dependent on speech rate, too? Xu's (1998) experimental results are instructive in that a coherent rising contour in Mandarin Chinese is being maintained across various speech rates: "the rising contour *as a whole* shifts more into the later portion of the syllable without systematic changes in the slope of the rise as the syllable duration increases" (italics mine). Notice that Xu uses disyllabic 'Rise-High' tone sequences, rather than the 'High-Rise-any lexical tone' context in (21)a, suggesting that slope identity may be suppressed by contextual tonal coarticulation (e.g. in "*conflicting*" context: H.LH, where two across-syllable tonal targets are unlike) but the loss of slope identity does not lead to absolute neutralization. On the other hand, slope identity is respected in "*compatible*" context (e.g. Rise-High: LH.HL, where two across-syllable tonal targets are like) at the same speaking rate (i.e. fast speed). Since slope is an essential dimension in making tonal contrasts, I assume that at least for those under the rubric of syllable tone languages, slope mappings have the following traits:

- Slope should be faithfully realized across various speech-rates if the neighboring tones are like tones.
- Slope may be flattened due to lack of time at fast speaking rate in "conflicting" context (see also Kuo et al. in press).
- Slope does not become steeper one in response to diminished duration.

Finally, Xu and Wang (2001) take Xu's (1998) finding as evidence for the existence of a linear (dynamic) pitch target. That is to say, a coherent rising contour in Mandarin Chinese is being maintained across various speech rates, indicating that rising tone is implemented to achieve a 'rising' target. In light of slope correspondence, this is basically reinterpretable as the implementation of slope identity.[6]

5.2.4 Outline of the Analyses

The bulk of this chapter is devoted to providing empirical evidence in support of slope correspondence. In §5.3, I discuss the 'faithfulness to the marked' phenomenon in Hangzhou Chinese. I.e. the contour of a more marked steep rising tone is faithfully rendered in sandhi while a smooth rising tone undergoes complete contour reduction in the same context. In §5.4, I will address the important role of the constraint against smooth-to-steep correspondent slopes in Lhasa Tibetan tone sandhi. I will show that the tone sandhi data in the above two languages are best characterized by slope correspondence. As a further step, slope identity preservation across syllables is dealt with in §5.5. The phenomenon in question is what I termed 'bounded tone extension.' I.e. peak delay in Mandarin Chinese and tone spreading in Wu Chinese are all subject to a 'locality' constraint. These 'extended' tones may move from their underlyingly affiliated syllables to the immediately adjacent toneless syllable, but never further to the second next toneless syllable. This locality constraint, as I will argue, is another instantiation of slope

---

[6] This is not to say that linear pitch targets do not exist.

correspondence. Section 5.6 concludes this chapter with a brief discussion of similar ideas pursued here in the former literature.

5.3 MATCH-SLOPE in Action: Hangzhou Chinese

Our first case study is concerned with the rising tones in Hangzhou Chinese (henceforth HZ), a dialect of Northern Wu Chinese (Qian 1992, S. Bao 2003). This language has four rising tones, three non-checked and one checked. The central goal of this section is to understand the asymmetrical behaviors of the two L-register non-checked rising tones in sandhi context (call them 'steep rise' and 'smooth rise'). The phenomenon in question is summarized as follows.

(22) Retention of the 'marked' tone in Hangzhou Chinese

Citation		Initial sandhi position
L-register steep rising tone (13)	→	H-register steep rising tone (13)
L-register *smooth rising tone (23)*	→	L-register *level tone (22)*

The contour of the steep rise is retained in initial sandhi position but the smooth rise is completely flattened in the same environment. It appears that the more marked rising tone, rather than the less rising marked tone, resists contour flattening. On the other hand, according to the conventional autosegmental representation, rising tones are regarded as the concatenation of L and H. Faithfulness constraints couched in unit correspondence predict that both rising tones should pattern alike (note further that they are both L-registered, which constitutes yet another challenge to the binary feature-based representational system). The issues raised in the brief discussion serve as the starting point of our analysis. To set the stage, some essential acoustic properties are provided in the subsequent section.

5.3.1 The Citation Tone Inventory

HZ has seven tones, including five (5) long tones and two (2) checked tones. To vindicate the former impressionistic transcriptions, I present the results from two separate phonetic studies in Figure 5-1. The plot on the left panel is based on my own acoustic measurements of the tokens extracted from the recordings in S. Bao's (2003) *The Phonetic Database of Hangzhou Chinese*. Normalized F_0 values and duration of the tokens (N=10 for each tone) were obtained using Yi Xu's Praat script (_TimenormalizeF0.praat, version 2.5.1).[7] On the right panel are the pitch contours replotted from Huang's (2001) experimental results, another published acoustic study of the HZ tonal system (N=15 for each tone).

(23) The inventory of Hangzhou Chinese citation tones

	My study	Huang (2001) (N=15)	Remarks
Tone 1	33	33	
Tone 2	23	23	
Tone 3	51	443	
Tone 4	455$^{!}$	334	
Tone 5	13	13	
Tone 6	44	44	Not included in Figure 5-1
Tone 7	23	23	Not included in Figure 5-1

---

[7] Available at http://www.phon.ucl.ac.uk/home/yi/downloads.html

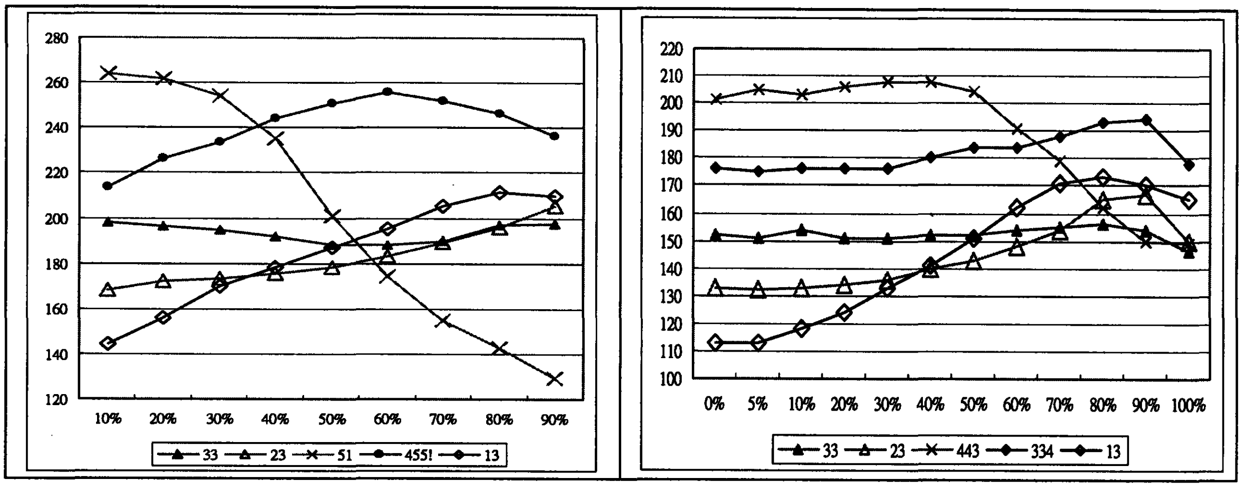


Figure 5-1 Normalized F_0 for citation non-checked tones in Hangzhou Chinese

As seen, the results of the two studies are largely corroborative, except for Tone 4.[8] According to the data reported here (on the left panel), Tone 4 ($455^!$) looks like an early rising tone: the F_0 curve starts at the midrange, reaches the F_0 maximum in the middle and falls towards the end. Here I use a raised exclamation mark to indicate this gentle fall. On the right panel, Huang's (2001) Tone 4 (334) can be described as a level tone with a small rise towards the end (or a *delayed* rise). In addition to this, there are substantial discrepancies between my data and Huang (2001) with respect to disyllabic tone sandhi. For the sake of simplicity, all the HZ tone sandhi data are based on my own study.

There is only one level tone in the non-checked citation tone inventory: Tone 1 (33). As we can see, its pitch fluctuation is in a range of around 10 Hz. Level tones in HZ can be accordingly defined: ΔF_0 (i.e. the F_0 difference between the two tonal targets) does not exceed 10 Hz. Among the three rising tones, Tones 13 and 23 are separated at the 10% time-point by 22.3 Hz (= 163.3 Hz – 141 Hz). Notice also that the two rising tones are both of the L-register; more precisely, 13 (*Yang Qu*) and 23 (*Yang Ping*) only occur with breathy voice-affected vowels (recall that HZ is

---

[8] Incidentally, it is noteworthy that the falling tone on the left hand side is of the continuous fall whereas the falling tone on the right hand side is clearly a delayed fall.

also a variety of Wu Chinese). So I take a 20 Hz tonal interval as a contrastive tone level (in the L-register). This completes our description of the phonetic attributes of the citation tones. In the following section, I turn to illustrate and discuss how citation tones are realized in word-initial sandhi position.

5.3.2 Phonetic Realization of Tones in Initial Sandhi Position

The core issue of this subsection is concerned with the phonetic realization of the sandhi tone inventory. All the tokens are extracted from S. Bao's (2003) recordings (N=10 for each tone). Like Shaoxing Chinese in chapter five, the HZ tone sandhi patterns in disyllabic lexical compounds are of the initial dominant type, which is presumably motivated by neutralization of murmur in non-initial position. The underlying contours of the initial tones are mostly retained in sandhi. The final tones are neutralized in sandhi contexts. One exception is that when two falling tones are in a row, i.e. 51-51, the output sequence is 44-31. This can be treated as a special tone sandhi rule. A comprehensive analysis of the full array of the relevant data is beyond the scope of this section. But the generalization below should suffice.

(24) Tonal alternations in initial sandhi position in HZ[9]

	Citation		Initial	Remarks
Tone 2	**23**	→	**22**	represented with 23' in Figure 5-2
Tone 5	**13**	→	**13**	represented with 13' in Figure 5-2
Tone 4	455$^{!}$	→	45	represented with 455!' in Figure 5-2
Tone 1	33	→	33	slightly raised by around 10 Hz
Tone 3	51	→	42	

---

[9] There are two checked tones in HZ, high level checked 44 and low rising checked 23. My observation is that both of them undergo complete contour reduction in initial sandhi position. This is because checked tones are in general shortened to around 100 ms in sandhi. Rising tones are impossible for such duration.

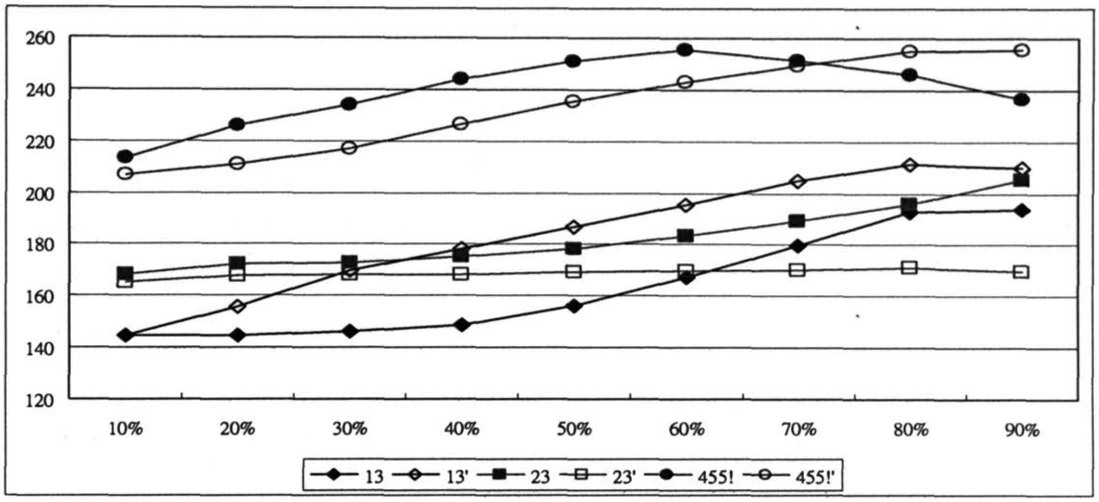


Figure 5-2 Normalized F_0 (in Hz) and duration for citation (unmarked) and sandhi forms (marked with an apostrophe) for the two rising tones in Hangzhou Chinese

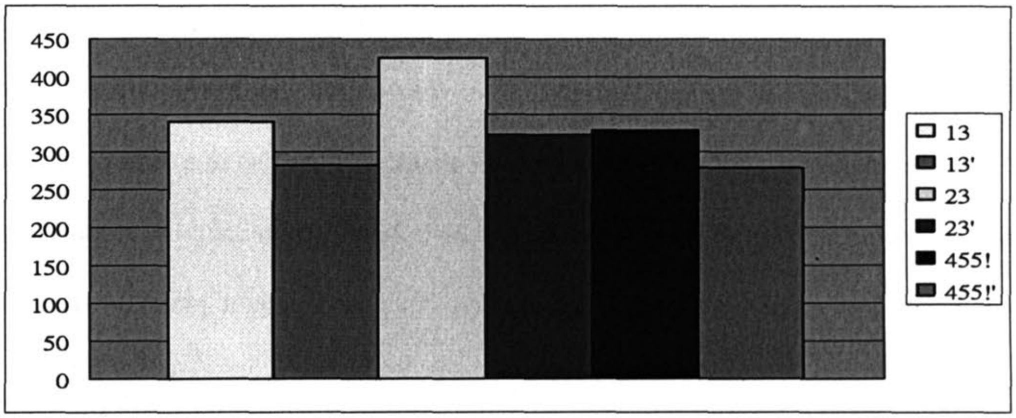


Figure 5-3 Rime duration (in ms) for the three rising tones in citation (unmarked) and initial sandhi forms (marked with an apostrophe)

From (24), it is evident that most citation tones are faithfully realized in initial sandhi position. As evidenced in the pitch tracks in Figure 5-2, however, a glaring anomaly emerges. The smooth rise 23 is completely flattened but the contour shapes of 13 and $455^!$ are retained: they become 13 and 45, respectively. Notice that for expository reasons, I use **23'** (whose actual phonetic form is 22) to represent the initial sandhi form of 23 and **13'/455!'** is used for the initial sandhi forms of 13/$455^!$ (whose surface forms are 13/45, respectively). The durational differences among the rising tones in citation and sandhi are illustrated in Figure 5-3. One thing worth mentioning is that tone 23 is longer than tone 13 by 50-70 ms. This distinction could serve as an additional cue for tonal discrimination.[10]

With the data in mind, our analysis proceeds as follows. In the first part, let us look at the sandhi forms of the two L-register rising tones: Tone 3 (23) and Tone 5 (13). Discussion of Tone 5 ($455^{!}$) is postponed to §5.3.4.

5.3.3 Faithfulness to the Marked as Slope Correspondence

Among the rising tones in HZ, the two L-register rising tones, Tone 2 (23) and Tone 5 (13), share certain essential affinities: both of them are L-registered, end at a similar F_0 point in isolation and belong to the continuous rise type. So it is puzzling why their surface forms are drastically different in contour shape in the same environment: word-initial sandhi position. More precisely, it is the more 'marked' tone 13, rather than tone 23, that keeps the rise in a diminished duration. I argue that the present 'faithfulness to the marked' phenomenon can be captured by slope correspondence. To begin the analysis, let us first consider relevant phonetic data and slope values below.[11]

(25) Tones 13 and 23 in citation and in initial sandhi position

	Tone		$\Delta F_0$10%-90%	Rime duration	Slope
L-register	Citation	13	65.1 Hz	341.2 ms	0.19
Sharp rise	Initial	13'	49.5 Hz	282.5 ms	0.17
L-register	Citation	23	36.9 Hz	425.9 ms	0.08
Smooth rise	Initial	**22**	4.6 Hz	323.4 ms	0.01

---

[10] On the other hand, the fact that 23 is longer than 13 indicates that a greater pitch excursion does not necessarily mean that a longer duration is needed. 23 may be awarded extra length for the sake of distinctiveness.

[11] Note that the in order to avoid onset and offset perturbation, F_0 difference is obtained between 10% and 90% of the normalized duration. However, rime duration is the length from 0% to 100%. This means that the slopes given in (25) are slightly shallower. For example, the actual slope of citation 13 is 0.23 (=65.1 Hz / (341.2ms*80%)), rather than 0.19, as shown in the first row in (25). This tiny difference can be ignored, because the difference between 0.23 and 0.19 would not be significant: ISD = log(0.23/0.19) = 0.07.

As shown, the steepness of the sharp rise is near-perfectly maintained in initial sandhi position. Recall that the Index of Slope Difference for a perfect slope match is 0 (= log1). For the case at hand, the ISD is 0.04 (= log(1.1) = log(0.19/0.17)). Provided that MATCH-SLOPE$_{\text{ISD}\rightarrow 0}$ is top-ranked, a sharp rise is licensed in a phonetically shorter syllable. For the present purpose, it suffices to employ the all-purpose tonal markedness constraint *RISE ('No rising tone is allowed in the output') in the analysis. *RISE must be ranked below the slope correspondence constraint MATCH-SLOPE$_{\text{ISD}\rightarrow 0}$. For convenience, the slope correspondence constraints are repeated below.

(26) MATCH-SLOPE$_{\text{ISD}\rightarrow 0}$

The slope of Tone *i* is m_i; the slope of Tone *j* is m_j, the slope of Tone K is m_k.
Tone *j* is the correspondent of Tone *i*; Tone *k* is the correspondent of Tone *i*.
Let the Index of Slope Difference of m_i-m_j be *p*.
Let the Index of Slope Difference of m_i-m_k be *q*.
Don't map Tone *i* to Tone *k* if $q>p>0$.

(27) MATCH-SLOPE$_{\text{ISD}>0}$ (i.e. 'NO SMOOTH-TO-STEEP CORRESPONDENT SLOPES')

The slope of Tone *i* is m_i; the slope of Tone *j* is m_j.
Tone *j* is the correspondent of Tone *i*.
Let the Index of Slope Difference of m_i-m_j be *p*.
Don't map Tone *i* to Tone *j* if $p<0$.

(28) Sharp rise in sandhi is licensed by high-ranked MATCH-SLOPE

Citation: 13 ΔF_0=65.1 Hz, 341.2ms	MATCH-SLOPE$_{ISD>0}$	MATCH-SLOPE$_{ISD\rightarrow 0}$	*RISE	
☞ a. 13 ΔF_0=49.5 Hz, 282.5ms			*	ISD=0.04
b. 14 ΔF_0=65.1 Hz, 282.5ms	*!	*	*	ISD=*-0.08*
c. 13$^{!}$ ΔF_0=39.5 Hz, 282.5ms		*!	*	ISD=0.17
d. 12 ΔF_0=29.5 Hz, 282.5ms		*!	*	ISD=0.27
e. 11 ΔF_0=10 Hz, 282.5ms		*!		ISD=1.27

Sample calculation of Index of Slope Difference (ISD):

$$\text{The ISD of Candidate (a)} = \log\left(\frac{\text{Citation Slope}}{\text{Sandhi Slope}}\right) = \log\left(\frac{\frac{65.1\,\text{Hz}}{341.2\,\text{ms}}}{\frac{49.5\,\text{Hz}}{282.5\,\text{ms}}}\right) =$$

$$\log\left(\frac{0.19}{0.17}\right) = \log(1.12) = 0.04$$

Candidate (b) has the same pitch shift as the citation form, 65.1 Hz. But this is a smooth-to-steep slope mapping (i.e. the ISD is negative). Candidate (c) is a reduced rise: let us hypothesize that its pitch excursion is 10 Hz less than that of the actual output (= 39.5 Hz = 49.5 Hz – 10 Hz). Candidate (d) is a further reduced rise whose pitch excursion is 20 Hz less than that of candidate (a). Candidate (e) is a level tone (Recall that 10 Hz is the upper limit 'tolerance range' for level tones). As seen, candidate (a) is selected as the optimal output because its ISD is closest to zero (0) among the candidates. The correspondent slopes in this candidate are basically identical in degree of steepness. Given that the all-purpose tonal markedness constraint *RISE is dominated, MATCH-SLOPE dictates the complete slope identity in the base-derivative relation (i.e. citation-sandhi).

As for citation tone 23, we have seen that its sandhi form is reduced to a L-register level tone. Given the present ranking, it is predicted that a gentle rise such as candidate (b) below would be selected as the winner. Consider now the following tableau.

(29) Why a perfect match in slope does not survive?

Citation: 23 ΔF_0=36.9 Hz, 425 ms	MATCH-SLOPE$_{ISD>0}$	MATCH-SLOPE$_{ISD\rightarrow 0}$	*RISE	
a. 23 ΔF_0=36.9 Hz, 323 ms	*!		*	ISD=*-0.1*
!☞ b. 23' ΔF_0=25.8 Hz, 323 ms			*	ISD≈0
☹ c. 22 ΔF_0= 4.6 Hz, 323 ms		*!		ISD=0.9

Some discussion is in order. For candidate (a), the pitch excursion is that of citation form. This output is ruled out because of the negative ISD. Recall from §5.2.3.1 that a negative ISD means a slope mapping from smooth to steep. So candidate (a) loses out because of the smooth-to-steep slope mappings. Importantly, minimization of articulatory effort is not at issue here. The reason is clear. It has been demonstrated in candidate (28)a that a 282.5 ms-long rime can bear a steep rise 13 (ΔF_0=49.5 Hz). Since the rime duration for tone 23 in sandhi is 323.4 ms (which is longer than the sharp rise 13 by **40.9** ms), all else being equal, there is no apparent reason as to why a less pronounced rising tone 23 cannot occur on a much longer rime. So I conclude that a more plausible explanation hinges on the bias against *smooth-to-steep* slope mappings.

Consider now candidate (b). The slope of citation 23 is 0.08. So a pitch shift of 25.8 Hz (=0.08*323.4ms) yields the same slope value in citation. The ISD for this mapping is extremely close to zero (0). In other words, this candidate should be selected as the winner. However, as we have seen, $23^!$ is *not* the actual surface form. Tone 23 changes to a level tone 22 in initial sandhi position, i.e. candidate (c). We then want to ask: why and how the requirement of slope similarity is voided for the case at hand? Let us consider some possible solutions.

First, avoidance of neutralization with tones in the sandhi tone inventory does not seem to work here. There are three L-register tones: i) the sharp rise 13 in sandhi is also separated by around 25 Hz in the tonal onset (see Figure 5-2); ii) as for the mid level tone 33, its sandhi form is slightly raised by 10 Hz and the pitch contour is positioned at around 210 Hz throughout. In other words, $23^!$ should have sufficient contrasts with its neighbors in the sandhi inventory.

(30) The contrasting tones of the hypothetical $23^!$ in the sandhi tone inventory

L-register	13, 33
H-register	45, 42

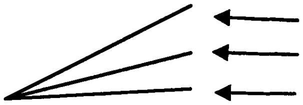
Notably, it is not possible to say that 22 and $23^!$ have an insufficient contrast (e.g. a constraint like *22-$23^!$), hence are neutralized. This is because 22 and $23^!$ are both the potential sandhi outputs for 23 in isolation. Even if 22 and $23^!$ are neutralized due to an insufficient contrast, $23^!$ still fares better by virtue of high-ranking slope correspondence constraint. Therefore, I conclude that tonal distinctiveness does not seem to at work for the case at hand.

Judging from the fact that the smallest rise in pitch is Tone 2 (23) in isolation (ΔF_0=35.6 Hz),[12] we can say that a 25.6 Hz difference is not "enough" for a rising tone in HZ. Under this view, the following illustration presents the gist of the idea pursued here.

---

[12] One may wonder the case of the checked rising tone, Tone 8 (23). Its ΔF_0 is 36.4 Hz, which is still greater than 35.6 Hz.

(31) Some potential output forms for citation 23 (cf. (29))



23: ISD<0: Slope mapping from smooth to steep
$23^{!}$: ISD≈0: Complete slope identity but insufficient ΔF_0
22: Deviant slope identity but it's a 'good' level tone

We have seen that the sandhi form 23 loses out because of smooth-to-steep mappings. The second potential output $23^{!}$ features complete slope identity preservation but $23^{!}$ also creates insufficient *syntagmatic* distinctiveness (i.e. candidate (29)b). By syntagmatic distinctiveness, I mean a sufficient contrast between the two tonal targets of a contour tone. As mentioned earlier, ΔF_0 of 25.6 Hz may not be "enough" for a rising tone in HZ. This intuition may be termed the 'no outlier condition.' We have assumed that ΔF_0 is insufficient for $23^{!}$ (NB: the lower limit is 35 Hz in HZ). It is not allowed to surface as 23 or a more pronounced rise, due to the disfavored smooth-to-steep slope mapping. So $23^{!}$ becomes an 'outlier,' whose pitch excursion is not enough for a rising tone and at the same time is too excessive for a level tone. As we have discussed earlier, the pitch fluctuation of level tones in HZ is within a range of 10 Hz. Note further that we are talking about F_0 intervals in the same pitch range: the discussion is solely concerned with tone 23 and the phonetic realization of its derivatives. So the 'unevenly-spaced F_0 interval' explanation is not relevant, either. Consider now the following illustration.

(32) $23^{!}$ as an 'outlier'

23 well-formed rise
$23^{!}$ *outlier*: not-so-well-formed rise/level tone
22 well-formed level tone

The guiding idea is that sufficient syntagmatic contrasts are necessary for a contour tone. So if an output form does not have a sufficient size of F_0 change, minimization of articulatory effort becomes operative. In a way, the present discussion is in line with the categorical nature of the phonemic inventory. As we have discussed, the contrasting tones $23^{!}$-22 may be neutralized but $23^{!}$ will be wrongly selected as the winner by the slope correspondence constraint. To void slope identity preservation, I resort to relational markedness (cf. Riggle 1999).

(33) EVEN-$\Delta(t_i\text{-}t_j)<10$ Hz (abbr. EVEN-$\Delta F_0<10$ Hz)

Tone T has two tonal targets, t_i and t_j, where t_i precedes t_j.
If t_i is p Hz, then t_j is $p\pm10$ Hz.

Relational markedness requires specific relation between two successive elements in the output. Under this approach, contour leveling is motivated by the requirement according to which ΔF_0 between the initial and the final target cannot exceed 10 Hz. In light of relational markedness, the analysis is provided below. We see that candidate (a) loses out because of an excessive pitch excursion for level tones. Likewise, candidate (b), a tiny rise is also banned. So candidate (c) is selected as the winner. Since 10 Hz is the 'tolerance range' for level tones, I assume that ΔF_0 for candidate (c) can range from 0 Hz to 10 Hz. Note also that the contour correspondence constraint *RELCORR*$(x<y)_{NUC}$ *is violated in (34)a-c* because these candidates are *not* (phonological) rising tones. I.e. the initial F_0 (x) and the final F_0 (y) do not stand in the 'smaller than' relation because ΔF_0 are all smaller than 35 Hz in (34)a-c. As a reminder, 35 Hz is the ΔF_0 lower limit for rising tones in HZ.

(34) Contour reduction revisited

Citation: 23	MATCH-SLOPE$_I$ SD>0	RELCORR $(x<y)_{NUC}$	EVEN $\Delta F_0<10$Hz	MATCH-SLOPE$_I$ SD→0
a. 23$^!$ ΔF_0=25 Hz		*	*!	
b. 22 ΔF_0=15 Hz		*	*!	*
☞ c. 22 ΔF_0=5 Hz		*		*
d. 23 ΔF_0= 36.9 Hz	*!		*	*

The reason why 23$^!$ does not surface is not because it is more marked than 22. If so, we expect that the steep rise 13 would be reduced to some extent or even undergo complete contour reduction, too. The empirical fact is that slope identity is well respected in the sandhi output of the steep rise 13. So it is reasonable to assume that slope identity should also be strictly enforced for the smooth rise 23. That being the case, contour leveling, or the loss of slope identity, is attributable to the condition according to which there is a clear-cut distinction between rising tones and level tones. Preservation of slope identity leads to a not-so-well-formed rising tone which incurs violation of the relational markedness constraint in (33); as a result, the contour of 23 is completely flattened.

Finally, it is important to note that rising tones are still possible in the inventory provided that RELCORR$(x<y)_{NUC}$ outranks the relational markedness constraint EVEN-$\Delta F_0<10$ Hz. As long as the 'smaller than' relation is established, this particular relation should be maintained in the output/sandhi, as the following tableau illustrates.

(35) 'Well-formed' rising tones are licensed by contour correspondence

Citation: 13 ΔF_0=65.1 Hz, 341.2ms	RelCorr $(x<y)_{\text{Nuc}}$	Even ΔF_0<10Hz	Match-Slope$_{\text{ISD}\to 0}$	
☞ a. 13 ΔF_0=49.5 Hz, 282.5ms		*		ISD=0.04
b. $13^{!}$ ΔF_0=39.5 Hz, 282.5ms		*	*!	ISD=0.17
c. 11 ΔF_0=5 Hz, 282.5ms	*!		*	ISD=1.27

From (35), it can be seen that the steep rise 13 surfaces as 13 in sandhi given the present ranking of constraints and at the same time, the contour leveling is induced if and only if the rise in pitch is not sufficient, e.g. the winning candidate (c) in (34). In sum, the difference in behavior between sharp and smooth rising tones is a consequence of the interactions between slope-matching and relational markedness.

This section is closed by reviewing an alternative analysis based on Zhang's (2002) Preserve-(tone). Since $23^{!}$ is more marked than 22 (i.e. well-established markedness ranking of tones: *Rise » *Fall » *Level, Yip 2002: 27-30), 22 will be chosen as the winning candidate by virtue of markedness. Consider now the analysis illustrated in the following tableaux. The constraints are basically equivalent to Zhang's (2002) formalism. *SteepRise should always dominate *SmoothRise, and *SmoothRise in turn dominates *UltraSmoothRise, due to the assumption according to which a steeper rising tone is more effortful, hence is more 'marked.' Let us further assume that duration is more or less fixed. I.e. as we have seen, citation forms are longer than sandhi forms, indicating that rime is not lengthened to accommodate a sharp rise. Finally, as a precursor of slope correspondence, Zhang's (2002) Preserve-(tone) constraints are given below.

(36) a. PRES(T, *i*): do not reduce 13 to 23
b. PRES(T, *I*): 13 must be faithfully realized.
c. PRES(T, *i*) » PRES(T, *I*)

The general schema for PRESERVE-(tone) is that a perceptually more deviant output is disfavored, according to Zhang's tonal similarity index. So an IO mapping 13 → 23 is worse than an IO mapping 13 → 13. This is expressed by way of the intrinsic ranking in (36)c.

Given this formulation, if PRESERVE-(tone) outranks tonal markedness constraints, it is predicted that the contour of the steep rise (13) is faithfully rendered, as illustrated in the following tableau.

(37) Why the steep rise is retained

Citation: 13	PRES(T, *i*)	PRES(T, *I*)	*STEEP RISE	*SMOOTH RISE	*ULTRASMOOTH RISE
☞ a. 13			*		
b. 23	*!			*	

(where steep rise=13, smooth rise=23.)

Likewise, for tone 23, the relevant PRESERVE-(tone) constraints are provided below. Since 22 is the most deviant output form for the input 23, mapping 23 to 22 incurs a fatal violation in the intrinsic ranking of constraints in (38)d. The analysis is given in the tableau in (39).

(38) a. PRES(T, *i*): do not reduce 23 to 22.
b. PRES(T, *j*): do not reduce 23 to $23^{!}$.
c. PRES(T, *I*): 23 must be faithfully realized.
d. PRES(T, *i*) » PRES(T, *j*) » PRES(T, *I*)

(39) Why the smooth rise is completely flattened

	Citation: 23	*SMOOTHRISE	*ULTRASMOOTHRISE	PRES (T, *i*)	PRES (T, *j*)	PRES (T, *1*)
	a. 23	*!				
	b. $23^{!}$		*!		*	*
☞	c. 22			*!		*

(where smooth rise=23, ultra-smooth rise=$23^{!}$.)

The total ranking is provided below. As shown, the present ranking of constraints predicts the expected output forms. Notice that PRESERVE-(13/23) are shorthand representations of the constraints in (36) and (38).

(40) PRESERVE-(13) » *STEEPRISE » *SMOOTHRISE » *ULTRASMOOTHRISE » PRESERVE-(23)

In terms of slope correspondence, this analysis amounts to the following: preservation of slope identity for the steep rise 13 is the most important; as a result, tonal markedness is overridden. In contrast, we have to say that slope identity of the less pronounced rising tone 23 is the least important; so minimization of articulatory effort takes effect and flattens out the entire contour (i.e. 23 → 22). Since Zhang's (2002) formalization of PRESERVE-(tone) incorporates magnitude of pitch movement, it is fair to describe the reasoning of the above analysis in terms of slope. This said, it is not clear why preservation of slope identity should be motivated in this way. Slope-matching should apply in a consistent manner, instead of to employ individual tone-specific slope identity preservation. I take the stance that slope correspondence is *not* relativized to different degrees of slope (i.e. steep rise vs. smooth rise). As we have discussed previously, it is not necessary to 'decompose' slope identity preservation in a tone-specific manner.

This completes our analysis of the two L-register rising tones in HZ. In the next section, I turn to the sandhi behavior of H-register rising tone $455^{!}$.

5.3.4 Slope Correspondence in the Early Rising Tones

For the H-register rising tone $455^{!}$, the F_0 maximum appears around the midpoint of the entire contour and the following F_0 curve falls towards the endpoint (see the figure in (42) below). As a first approximation, $455^{!}$'s contour shape can be regarded as the early rise type. As I have assumed at the outset of this chapter, the size of pitch excursion between the two slices of Tone T serves as the altitude change dimension of slope. This assumption is based on the continuous rise/fall contour tones, i.e. an idealized *straight* pitch contour. As a further step, the goal of this section is to look at the following issue: how does slope correspondence work in non-continuous contour tones, in particular, an early rise (note again that this is not a straight rise)? Consider now the relevant data below.

(41) Slope-matching for H-register rising tone in Hangzhou Chinese

a. From smooth to steep

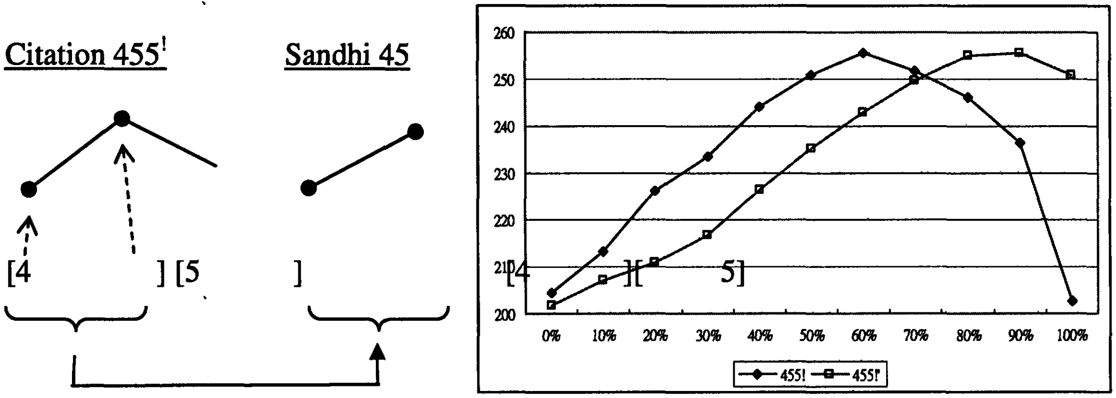
Context		$\Delta F_0$10%-90%	Rime duration	Slope
Citation	$455^{!}$	23 Hz	328.3 ms	0.07
Initial	45	48.5 Hz	279.1 ms	0.17

b. From steep to smooth

Context		$\Delta F_0$10%-60%	Rime duration	Slope
Citation	$455^{!}$	42.2 Hz	164.1 ms (=328*50%)	0.25
		$\Delta F_0$10%-90%	Rime duration	
Initial	45	48.5 Hz	279.1 ms	0.17

From (41)a, it is obvious that the F_0 shift size in citation is smaller than that in sandhi, if taken from 10%-90% of normalized duration. Alternatively, if we take the F_0 change between the minimum (10%) and the maximum (60%), as in (41)b, we find a similar slope-matching pattern to the preceding section. The idea is illustrated below.

(42) Tone 455^1: slope correspondence in early rise



From the above pitch tracings, it can be seen that the F$_0$ curve corresponding to the citation tone 455^1 drops after reaching its F$_0$ maximum of the 60% time-point. It is reasonable to say that only the slope value of the first 60% part is used for slope similarity comparison. It suggests that slope can be obtained between the F$_0$ minimum and the F$_0$ maximum, graphically represented on the left hand side of (42). As we can see, the second half of 455^1 drops to the midrange, especially between the 90% and 100% time-points. In contrast, 455^1's sandhi form, represented with **455!'** in the above figure, does not feature the same contour shape at the endpoint. Instead, **455!'** looks like a continuous rising tone. I take this as the evidence indicating that there is no phonological specification on the *right* edge of the citation 455^1.

In sum, the HZ data seem to suggest that slope correspondence may be sensitive to only some portion of a pitch contour, in particular, when the F$_0$ maximum or the F$_0$ minimum is not properly aligned with the edges.

5.3.5 Summary of this Section

To recapitulate the foregoing discussion, we began with the puzzling phenomenon whereby all else being equal, a sharp rise is preserved but a smooth rise is completely flattened in initial sandhi position. As we have seen, duration and adjacent tones (i.e. the first two factors of contour tone licensing in (1)) do not come into play for the present case. More precisely, duration is equally diminished in initial sandhi position and non-initial tones do not influence the phonetic form of the initial tone. With regard to faithfulness (i.e. unit correspondence), this is unpredicted because the two L-register rising tones have the same tonal melody, LH (modulo the representational issue), indicating that they should pattern alike. But this prediction is not borne out. In view of tonal markedness, steep rising tones should be more markedly complex than smooth rising tones. In other words, *STEEPRISE should always outrank *SMOOTHRISE. But what we have seen is that steep rise is preferentially retained. As I have argued, this 'markedness reversal' phenomenon is better treated as correspondence. I have also demonstrated that slope correspondence (with relational markedness) provides a unified account for the 'faithfulness to the marked' puzzle.

For HZ, the discussion is mainly concerned with the issue of slope identity preservation (i.e. MATCH-SLOPE$_{\mathrm{ISD}\to 0}$). In the next case study of Lhasa Tibetan, let us turn to the role of MATCH-SLOPE$_{\mathrm{ISD}>0}$ in tone sandhi, i.e. the bias against *smooth-to-steep* slope mappings.

5.4 Lhasa Tibetan: Against Smooth-to-Steep Slope Mappings

We have discussed in §5.2.3.1 that the slope-matching mechanism also features the alleged dispreference for the mappings from smooth to steep. The goal of this section is to provide evidence and show that this dispreference plays a substantial role in licensing rising tones in the

Lhasa Tibetan (hereafter LT) tonal system.

As we will see shortly, LT has two basic tonal melodies: H vs. $\overset{\frown}{\mathrm{LH}}$ (Duanmu 1992) and a two-way contrast of rime length: long vs. short. For disyllabic lexical compounds, we would expect that there are sixteen (16=4*4) tone sandhi patterns. As a matter of fact, only three (3) patterns are attested.

(43) A gross generalization of Lhasa Tibetan disyllabic tone sandhi

Combination	Mora count	Attested patterns	Unattested Patterns
Short-Short	μ-μ	H-H/L-H	*H-LH/*LH-LH/*LH-H
Short-Long	μ-μμ	H-H/L-**$\overset{\frown}{\mathrm{LH}}$**	*H-LH/*LH-LH/*LH-H
Long-Long	μμ-μμ	H-H/ L-**$\overset{\frown}{\mathrm{LH}}$**	*H-LH/*LH-LH/*LH-H
Long-Short	μμ-μ	H-H/L-H	*H-LH/*LH-LH/*LH-H

Of particular interest is the rising tone on the word-final syllable (in boldface above). It seems to be the case that rising tones are licensed only if the second syllable is long. At first blush, this generalization is in accord with the cross-linguistically robust phenomenon of phrase-final lengthening. It has been pointed out that phrase-final position, rather than phrase-initial position, is endowed with the privileged contour bearing ability (Zhang 2002, *inter alia*). Consequently, it is conceivable that rising tones are not licensed in initial sandhi position even though the initial syllable is long.

However, this cannot be the end of the story. In isolation, short syllables bear a rising tone, too. It is unlikely that a long syllable in initial sandhi position is shorter than a shorter syllable in isolation. This conjecture is confirmed by the phonetic data reported here (to which I shall return shortly): a word-initial long syllable is still longer than a short monosyllable by about **60** ms. Furthermore, the pitch excursions of the actual rising tones (both in isolation and sandhi) are largely similar, ranging from 30 to 40 Hz. Under the 'licensing-by-duration' approach (i.e. (1)i), if a short syllable in isolation can carry a rising tone, it is unexpected that a long syllable in initial sandhi position cannot do so, given that this long syllable in sandhi is still longer than the short syllable in isolation by 60 ms. So it is fair to say that a duration-based account cannot provide a reasonable account for the absence of contour tones on the word-initial long syllable. On the other hand, according to the conventional autosegmental representation, contour tone licensing is sensitive to mora count. That is, rising tones occur only on bimoraic syllables. It appears that this approach does not fare better for the case at hand because the word-initial long syllable should be bimoraic, too.

To rule out contour tones in initial sandhi position, one common diagnostic is to invoke the general-purpose positional markedness constraint (see Zhang 2002 for general argument against this type of constraints). For example, Yip (2002) proposes LICENCE-CONTOUR to penalize word-initial rising tones in Lhasa Tibetan.[13] Provided that LICENCE-CONTOUR is top-ranked, no contour tone survives in initial sandhi position. Thus, it seems to be the case that Lhasa Tibetan calls for the need of general-purpose positional markedness constraints. In the following section, I would like to argue that the 'no word-initial rising tone in sandhi' phenomenon should be treated as slope correspondence. In particular, the ban on smooth-to-steep slope mappings serves as a better motivation for the phenomenon in question.

5.4.1 Acoustic Properties of Monosyllabic Tones

It has been established that LT tones are customarily reduced to binary H- vs. L- register contrasts, with the surface pitch contours as a function of the rime shape. The conventional transcriptions are tabulated below. Tone values are given on a five-point scale. Notice that rime

---

[13] Zoll's (1998) COINCIDE constraint family fits the bill equally well.

type is my interpretation, based on the experimental results reported in this section.

(44) Surface tonal contours on monosyllables in Lhasa Tibetan (Hu 1980, Hu *et al.* 1982)

Rime type		H-register	L-register
Short	V	54	12
Long	V:/VN/ṽ:	55	113
		55	113
Glottalized	Vp̚/V̰	52	132
	VN̰/V̰:	52	132

It has been debated in the past few decades as to whether LT is a two-, four- or six-tone system (see Sun 1997 for an overview). As we can see, tones on the short and long rimes can be collapsed into one category; the contour shapes of short-rimed tones may be due to undershoot. The problematic case concerns the tones on what I termed 'glottalized rime.' In the literature, it has been claimed that high falling tone (52) and the convex tone (132) occur with a glottal stop coda. Diachronically speaking, these glottalized rimes are derived from the loss of coda obstruent (but note that –p does not drop) and the loss of –s in nasal-s coda clusters. Two exemplars of the glottalized rimes are illustrated below: a glottalized nasal [saŋ̰^{52}] on the left panel and a glottalized short vowel [p^{h}a̰^{12}] on the right panel (see below for the data source). As we can see, the final portion of rimes of this sort (the circled part) is glottalized: voicing pulses are further apart than they are in modal voicing. As Silverman (2002) points out, harmonic structures may not be well defined on non-modal phonation, resulting in poorly cued F_0. This is also evidenced in LT. Glottalization may break the F_0 contour, making an abrupt pitch jump around the midpoint of a rime (e.g. the pitch tracks on the left panel) or may eliminate the F_0

contour (e.g. the final 'missing' part of the contour on the right panel).

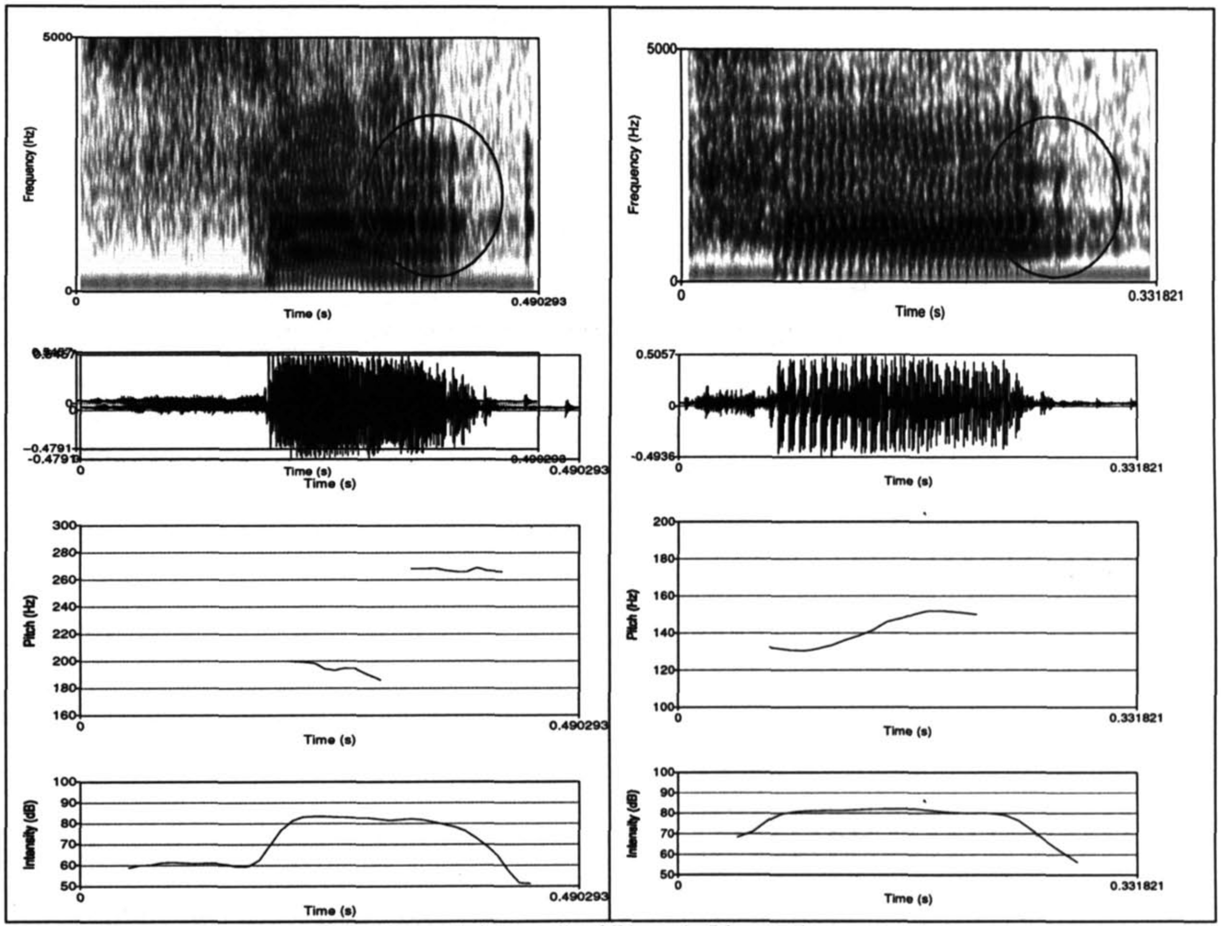


Figure 5-4 Glottalized Rimes in Lhasa Tibetan

It appears that the so-called high falling/convex tone in the descriptions of Tibeto-Burmanists is in actuality a matter of glottalization. Turning back to the tonemicization issue, we draw the conclusion that the high falling/convex tones can be treated as H- or L-register tones on non-modal phonation. I do not discuss the case of tones on the glottalized rimes since pitch tracks are not reliably obtained on them.

Let us now turn to observe the F_0 curves of LT tones. The phonetic data for Lhasa Tibetan were extracted from a cassette tape accompaning *A Pronunciation Guide for Lhasa Tibetan*

(Zhou 1983). It consists mainly of monosyllabic and disyllabic words read in isolation by one male speaker of Lhasa Tibetan. The acoustic measurements were made using praat with Yi Xu's script (_Timenormalize.praat). 25 tokens were measured for each tone.

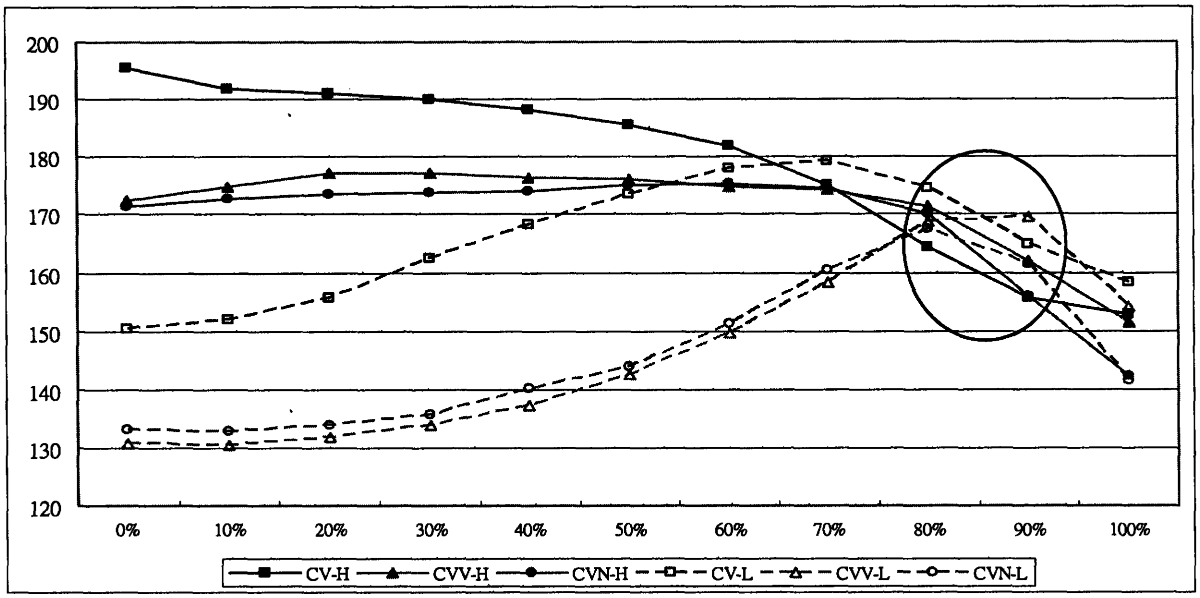


Figure 5-5 Normalized F_0 for Lhasa Tibetan Citation Tones

Where CV=short vowel, CVV=long vowel, CVN=vowel-nasal, H=H-register, L=L-register. As seen, there is an obvious distinction between the H-register tones (represented with filled shape) and the L-register tones (represented with hollow shape and dashed line). All L-register tones start somewhere below 160 Hz. By contrast, all tones move to a similar point at the 90% point: somewhere between 160 Hz and 170 Hz, as indicated by the circle. The final F_0 lowering in the 100% time-point is attributable to the accompanying glottalization when a syllable is being terminated (or perhaps the domain-final effect). To this end, we draw the conclusion that Lhasa Tibetan has a two-level tonal contrast and the cut-off point is 160 Hz. In addition to this, the F_0 curves corresponding short-rimed tones (i.e. CV-H and CV-L) are raised by around 20 Hz up to 50% of normalized duration in comparison to their counterparts on long rimes (i.e.

CVV-H/CVN-H and CVV-L/CVN-L). It is not clear to me why this should be the case. A plausible account is that it might be more difficult to begin with the precise H or L targets in durationally shorter syllables, so the starting point is raised.

The pitch tracings are instructive in that the LT level tones (i.e. H-register tone) and rising tones (i.e. L-register tone) can be defined as follows. I adopt Duanmu's (1992) proposal according to which LT has two underlying tonal melodies, H and L͡H. While the final F_0 drop can be attributed to glottalization or a boundary tone, the rising contours of the L-register are preserved when the condition is met (see below for more details). Given this assumption, as mentioned earlier, the pitch excursions for H- and L-register tones are implemented in a systematic fashion. 20 Hz is by and large the tolerable range of fluctuation for level tones while a pitch shift of around 30 Hz is the lower limit for rising tones.

(45) Pitch range for two tonal melodies of Lhasa Tibetan

Maximal pitch range for H-register tone (level)	< 20 Hz
Minimal pitch excursion for L-register tone (rising)	> 30 Hz

The F_0 values were in turn transformed into Logarithmic Z-score (LZ) and plotted in the following diagram. As we can see, the effect of LZ normalization is not so significant: the LZ score values roughly correspond to the normalized raw F_0 values: 1 in LZ is around 20 Hz. This is mainly due to the fact that LT contrasts only two tone levels. So I will use the normalized F_0 values throughout the discussion that follows.

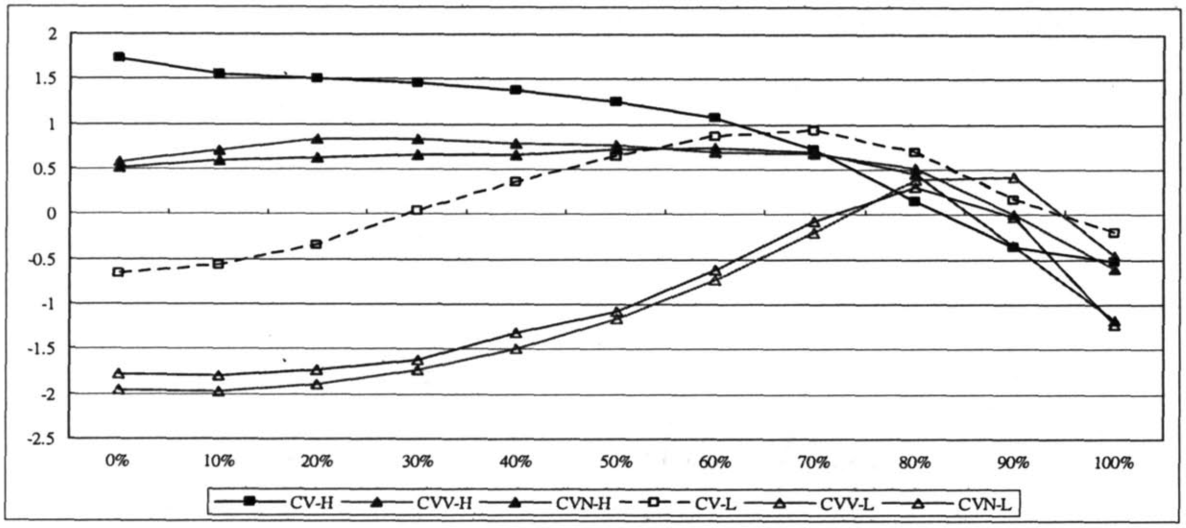


Figure 5-6 LZ normalized F_0 for citation tones in Lhasa Tibetan

The absolute duration and mean intensity for H- and L-register citation tones are presented below. Regarding duration, long rime (around 300 ms) is basically two times longer than short rime (around 150 ms).[14] Moreover, L-register tones are invariably longer than H-register tones. This seems to be a universal tendency of tone production (Maddieson 1978).

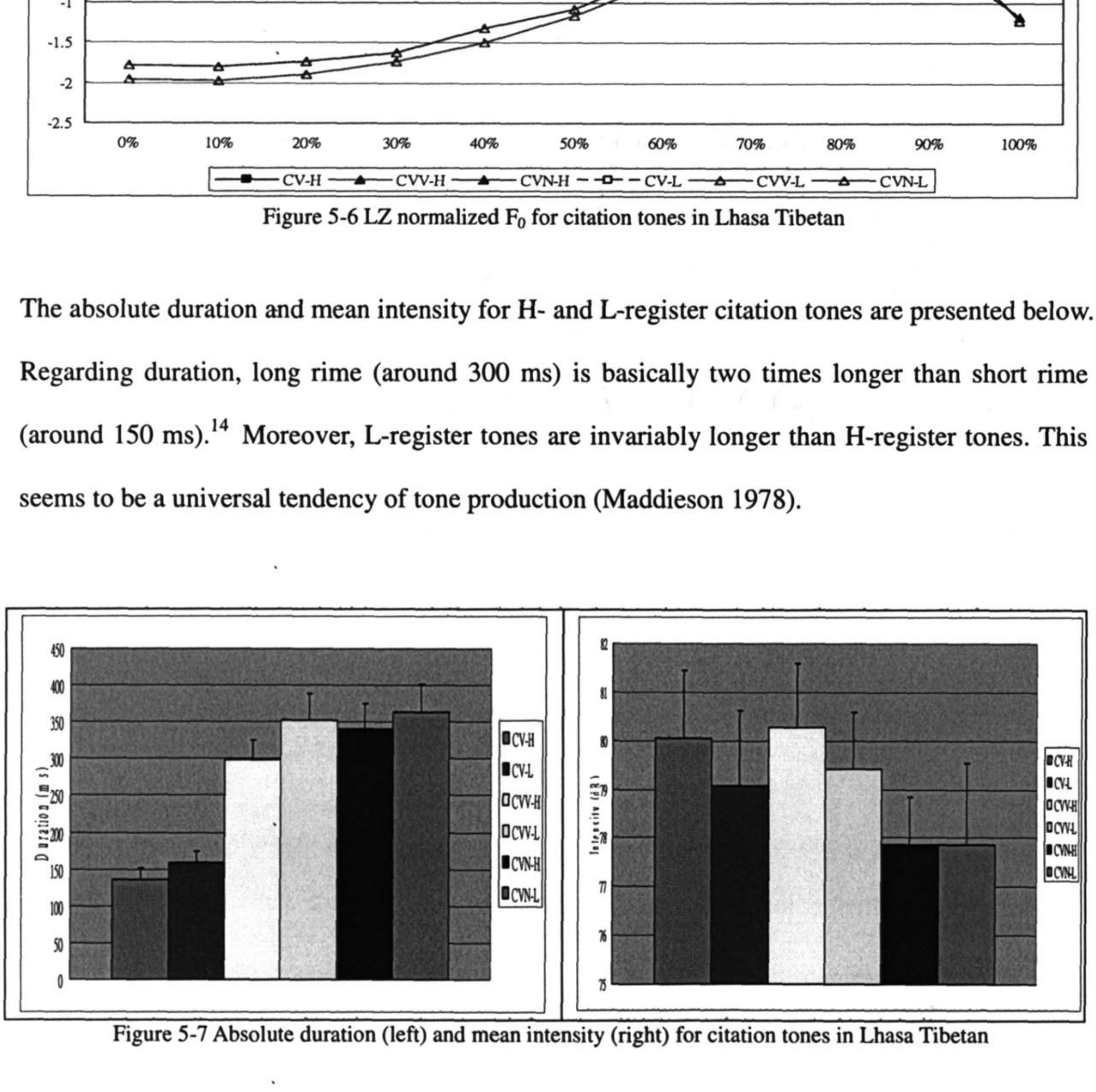


Figure 5-7 Absolute duration (left) and mean intensity (right) for citation tones in Lhasa Tibetan

---

[14] Notice that I do not include the data for the glottalized short vowel and the glottalized VN rime. According to my measurements, their average durations are about 100 ms and 200 ms respectively. As far as I can tell, the glottalized VN rime pattern on a par with short rime in tone sandhi, suggesting that it is *not* categorized as the long rime.

As for intensity, L-register tones are slightly less sonorant than H-register tones. Furthermore, tones on the VN rime are least sonorant, which is attributed to the presence of nasal coda.

The above is a brief sketch of acoustic properties of citation tones in LT. In particular, I have shown that LT has two tonal contrasts, H and LH and the rime length distinction is robust. In general, the phonetic data reported here corroborate the previous impressionistic or experimental studies and are essential in understanding the disyllabic tone sandhi patterns, to which I turn in §5.4.2.

5.4.2 Acoustic Properties of Disyllabic Tone Sandhi

We have mentioned at the outset that there are only three major patterns attested in disyllabic tone sandhi: H-H, L-H and L-L͡H. Tonal alternations for 8 combinations (=2 tones*2 length distinctions*2 positions) are presented below. For each combination, 10~20 tokens were measured. The second tones are unmarked because they are predictable (see below for details).

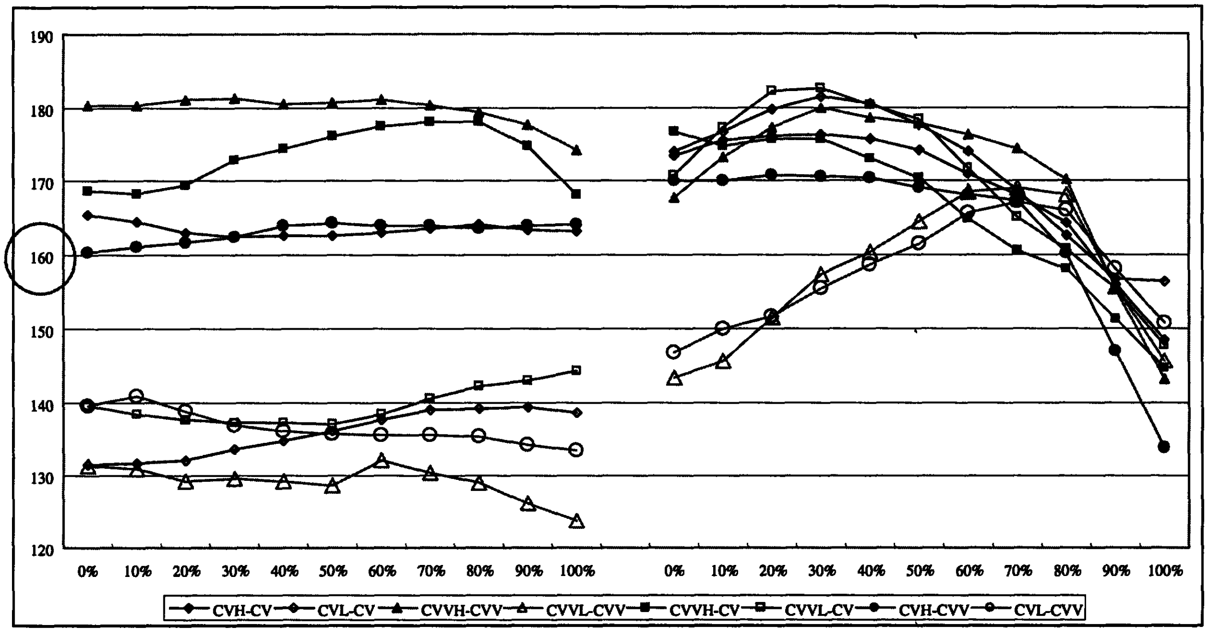


Figure 5-8 Normalized F_0 for the three disyllabic tone sandhi patterns in Lhasa Tibetan

The F_0 tracings of disyllabic tone sandhi patterns reveal the following facts. First, tones are completely flattened in initial sandhi position. This is in accord with previous impressionistic or instrumental studies. As seen, no initial sandhi tone has an overall F_0 fluctuation that is more than 20 Hz (cf. (45)).

Second, the H- and L-register contrast remains distinctive. The cutoff point is still 160 Hz throughout. By contrast, most sandhi final tones are neutralized to the H-register. In general, the F_0 curves on the right hand side of Figure 5-8 crowd in the higher part of the pitch range. This finding also corroborates with the previous impressionistic or instrumental studies of LT tonal system.

Third, the present data indicate that the rising tones in sandhi final position do not entirely belong to the H-register (contra Hu *et al.* 1982).[15] The corresponding F_0 curves are those of CVVL-CVV and CVL-CVV. The alleged H-register rising tone in sandhi final position has been transcribed as 24 in Hu *et al.* (1982), in contrast with its allotones 12/113 in isolation. From the diagram above, it should be obvious that the first 50% of the curves is positioned in the lower part of the pitch range, indicating that it is better to treat the sandhi rising tone as a composition of L and H.

Fourth, the F_0 curves converge towards the end of a disyllable, i.e. the 80%-100% time-point of the word-final tone. I assume that this phenomenon is induced by a boundary tone L-. Recall from our discussion on citation tone that a similar pattern is observed, too, albeit less pronounced.

The above discussion suffices for the present purpose. The other essential acoustic properties,

---

[15] This fact weakens some arguments made in Yip (1993) and Sun (1997). In particular, Yip (1993) argues that the laryngeal node is deleted and then is inserted with the default H-register because all tones are neutralized to H-register. Sun (1997) argues against Duanmu's (1992) analysis by mentioning that the rising tone in sandhi final position is a H-register rise, rather than LH (=L-register rise) per se. The present data confirm Duanmu's (1992) proposal.

in particular, rime duration and mean intensity, are described below. I discuss them in turn.

As far as rime duration is concerned, we see from the left panel of Figure 5-9 that the length distinction is systematically maintained. Phrase-final lengthening is weak. It is probably due to the fact that monosyllabic rimes in isolation are approximately two times longer than their counterparts in the disyllabic sandhi context. In other words, monosyllabic rimes are roughly as long as the two rimes in a disyllable. A comparison is given in (46).

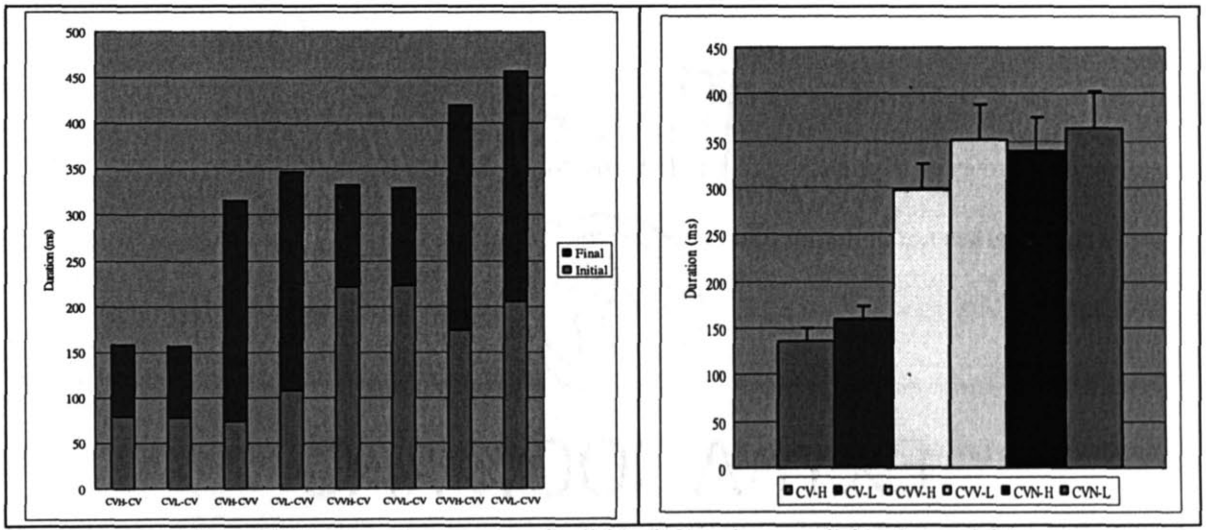


Figure 5-9 Duration for rimes in a disyllable (left) and monosyllabic rimes (right)

(46) Duration for monosyllabic rime and disyllabic rimes

Isolation		Disyllabic combination		Rime duration
Short	≈	Short-Short	≈	150 ms
Long	≈	Long-short	≈	300 ms
	≈	Short-Long	≈	300 ms
N/A		Long-Long		400~450 ms

As first blush, this finding weakens our discussion in the beginning according to which the rising tones appear only in sandhi final position simply because of the phrase-final lengthening effect. I would like to postpone the relevant discussion to §5.4.6 and show that phrase-final lengthening does play a role, although this effect is comparatively minor.

Finally, the data for mean intensity is provided below. It looks as if the final syllable consistently carries greater intensity than the initial syllable. This observation was confirmed by a t-test: p=0.001. That being the case, it is quite questionable if LT could be treated as a language with initial prominence in tone sandhi. As we have discussed, the initial prominence view is widely accepted because non-initial tones are neutralized. Initial prominence is supposed to feature some phonetic correlates, in particular, duration or intensity. Or plainly, initial syllables are stressed so as to motivate initial prominence. The data reported here show the reverse distribution: none of these properties for stress is privileged in initial syllables. Instead, word-final syllables are better characterized as being stressed.

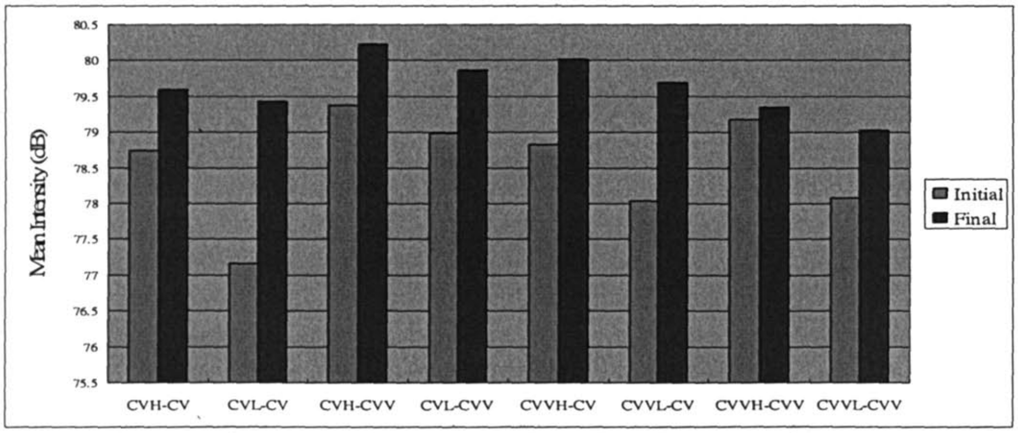


Figure 5-10 Mean intensity for Lhasa Tibetan disyllabic tone sandhi patterns

This completes the description of acoustic properties of the disyllabic tone sandhi in LT. In light of these findings, several analytical issues will be addressed in the following section.

5.4.3 Discussion and Analytical Issues

From the foregoing description of acoustic properties of the LT tonal system, we arrive at the following generalizations. First, there are two basic tonal melodies, H and LH. We may still refer the two tonal contrasts as H- and L-register tones. Second, three tone sandhi patterns are attested in a disyllabic domain: H-H, L-H and L-LH. Third, all initial tones are completely flattened regardless of the length distinction, while rising tones appear only if i) the initial syllable carries LH and ii) the final syllable is long. Fourth, word-final syllables are stressed, by virtue of the fact that word-final syllables are (slightly) longer and carry more intensity. With these points in mind, the core data are tabulated below. There are several interesting issues arising from the data. I list those central to the analysis in (48).

(47) Schematic Disyllabic Tone Sandhi Rules

Final Initial	Short-H	Short-LH	Long-H	Long-LH
Short-H	H-H	H-H	H-H	**H-H**
Short-LH	*L-H*	*L-H*	L-LH	L-LH
Long-H	H-H	H-H	H-H	**H-H**
Long-LH	*L-H*	*L-H*	L-LH	L-LH

(48) Some important analytical issues for LT disyllabic tone sandhi

a. Why most tones are neutralized to H in word-final syllables?
b. Why all LH's are not licensed in word-initial long syllables? (=the italicized)
c. Why some LH's are licensed in word-final long syllables? (=the shaded cell)
d. Why some LH's are *not* licensed in word-final syllables? (=the boldfaced)

The following sections are devoted to answer the questions raised above. Note that I do not discuss the H-H sequences (i.e. the unmarked cells in (47)). This is simply because these combinations do not involve substantial tonal alternations.

5.4.4 Neutralization of the Non-initial Tones

Our analysis begins with the following premise. LT tone sandhi is *not* motivated by initial prominence due to the fact that the initial sandhi position lacks phonetic correlates for stress. Instead, word-final syllables are stressed. To capture the fact that most tones are neutralized in sandhi final position, I revise de Lacy's (2002b) proposal of tone-stress interaction and propose that H-register tones are more compatible with stressed syllables and L-register tones are more likely to appear in unstressed syllables: we have learned from Figure 5-7 that H-register tones are more sonorous than L-register tones. The *HEAD/L-REGISTER constraint requires that L-register tones (LH) do not appear in the prosodic head, i.e. the stressed syllable, while the *HEAD/H-REGISTER constraint demands that H-register tones do not occur in the stressed syllable. Given these, the two constraints are intrinsically ranked as follows.

(49) *HEAD/L-REGISTER » *HEAD/H-REGISTER (abbr. *HD-H » *HD-L)

Conversely, H-register tones are disfavored in prosodically 'weak' position (or non-head, i.e. unstressed syllables), while L-register tones are more compatible with unstressed syllables. This tendency is captured by the following ranking of constraints. *NON-HEAD/H-REGISTER means 'no H-register tone is allowed in unstressed syllables' and *NON-HEAD/L-REGISTER is defined as 'no L-register tone is allowed in unstressed syllables.'

(50) *NON-HEAD/H-REGISTER » *NON-HEAD/L-REGISTER (abbr. *NON-HD-H » *NON-HD-L)

Let us first look at how these two sets of constraints work in sandhi initial position. The analysis is given in the following tableau.

(51) Word-initial tones

L-T	*HD/L	*HD/H	MAX-(H)	*NON-HD/H	*NON-HD/L
☞ a. L-T*					*
b. H-T*				*!	
H-T	*HD/L	*HD/H	MAX-(H)	*NON-HD/H	*NON-HD/L
☞ c. H-T*				*	
d. L-T*			*!		*

(Note that H=H-register, high tone, L=L-register, i.e. rising tone, the asterisk marks the stressed syllable, T=any tone)

From (51), we can see that if MAX-(H) outranks *NON-HD constraints, it is predicted that underlying tones are faithfully realized in initial sandhi position.

On the other hand, as we will see, the presence of L-register tones in stressed syllable (i.e. final position) is licensed by MATCH-SLOPE. To satisfy *HD/L, (52)b becomes a high level tone, but this candidate is ruled out by the slope correspondence constraint.

(52) The preservation of rising tones in stressed syllables

T-L	MATCH-SLOPE	*HD/L	*HD/H
☞ a. T-L*		*	
b. T-H*	*!		*

(Note that L=L-register, rising tone, H=H-register, high tone, T=any tone, *=stress)

In contrast, if an underlying rising tone violates MATCH-SLOPE on the surface (e.g. in case of the smooth-to-steep slope mapping), the current ranking of constraints predicts is that neutralization to the H-register takes place, for instance, Short^{T}-**Short**^{LH} → Short^{T}-**Short**^{H}.

(53) *L*: an output rising tone that violates MATCH-SLOPE

T-L	**SMOOTH-TO-STEEP SLOPE MAPPING*	*HD/L	*HD/H
a. T-*L**	*!	*	
☞ b. T-H*			*

(Note that L=L-register, rising tone, H=H-register, high tone, T=any tone, *=stress)

Recall the first question in (48)a, namely, why are most tones neutralized to H-register in final position? The core idea is that if an underlying rising tone (=L-register) cannot be faithfully rendered in stressed syllable, *HD/L » *HD/H predicts that a H-register tone (=high tone) will be chosen as the optimal output. In other words, the presence of word-final rising tones is arguably attributed to the high-ranked slope correspondence constraint.

With the above discussion in mind, I set out to explain how and why the rising tones are or are not licensed by slope correspondence in the following sections.

5.4.5 No Rising Tone as No *Smooth-to-Steep* Mapping

We have learned that the duration of the short syllables is around 80 ms in sandhi, a half of the rime duration in isolation. The phrase-final lengthening effect is weak. Let us simply assume that duration is more or less fixed in sandhi. I.e. in general rime lengthening is not used as a strategy to faithfully realize an underlying rise in LT. Let us first look at the case in which short LH becomes short H in final position. The analysis is illustrated below.

(54) $\text{Short}^{H}\text{-Short}^{LH} \rightarrow \text{Short}^{H}\text{-Short}^{H}$

Citation: <H_1>-LH_2 $\Delta F_0(LH_2)$=30 Hz Rime=160 ms	MATCH-SLOPE$_{ISD>0}$	MATCH-SLOPE$_{ISD \rightarrow 0}$	RELCORR $(x<y)_{Nuc}$	
a. <H_1>-LH_2 ΔF_0 (LH_2)=30Hz <79 ms>-80 ms	*!	*		ISD = *-0.31*
b. <H_1>-H_2 ΔF_0 (H_2)=20 Hz <79 ms>-80ms	*!	*		ISD = *-0.14*
☞ c. <H_1>-H_2 $\Delta F_0(H_2)$=15 Hz <79 ms>-80ms			*	ISD ≈ 0

(Note that word-initial tones in angle brackets are not analyzed.)

Candidate (a) keeps the size of pitch shift in citation, 30 Hz, resulting in a smooth-to-steep mapping. Candidate (b) incurs the same violation, hence is ruled out. Therefore, we expect that candidate (c) is chosen as the winner.[16] Notice also that H_1-L_2 is a possible winner but this output is ruled out by this ranking *HD/L » *HD/H, as discussed in (53). The realization of LH on word-initial syllables can be analyzed in the same fashion, hence is not included here.

So far we have discussed why LH does not occur in sandhi short syllables. As a matter of fact, tonal markedness also predicts that LH undergoes complete contour reduction in sandhi short syllables. This is because the duration of a short syllable, be it in initial or in final sandhi position, is generally shortened to about 80 ms. According to Xu and Sun's (2002) linear equation, 80 ms is not enough to implement a 30 Hz pitch rise (i.e. the lower limit of rising tones in LT). The minimum time for a rise is greater than 100 ms. As briefly mentioned at the outset, however, tonal markedness fails to account for the fact that LH is also unattested in initial long syllables. Long syllables have a mean duration of around 200 ms in initial position. This fact precludes a tonal markedness account because 200 ms is far more than enough for a 30 Hz pitch rise. In terms of slope-matching, realizing a 30 Hz pitch rise may lead to a smooth-to-steep mapping. As a result, LH is not licensed in long syllables in sandhi. The analysis is illustrated in the following tableau.[17] Note that in isolation the average pitch excursion of the rising tones in long syllables (around 40 Hz) is greater than that in short syllables (around 30 Hz). See also Figure 5-5.

---

[16] Notice that the actual pitch fluctuation is greater than 20 Hz (cf. Figure 5-8). I assume that this is attributable to the presence of the low boundary tone L-.

[17] It might well be the case that tone absorption takes place in the present case: Long^{LH}-Short^{H} → Long^{L}-Short^{H}. But the tone absorption account fails to predict why Long^{LH}-Long^{H} → Long^{L}-Long^{LH} (*Long^{L}-Long^{H}).

(55) Long^{LH}-Short^{H} → Long^{L}-Short^{H}

Citation: LH_1-<H_2> $\Delta F_0(LH_1)$=40 Hz Rime=352 ms	MATCH-SLOPE$_{ISD>0}$	MATCH-SLOPE$_{ISD\rightarrow 0}$	RELCORR $(x<y)_{Nuc}$	
a. LH_1-<H_2> $\Delta F_0 (LH_1)$ = 40 Hz 222 ms-<107 ms>	*!	*		ISD = 0.78
b. LH_1-<H_2> $\Delta F_0 (LH_1)$ = 30 Hz 222 ms-<107 ms>	*!			ISD = *-0.08*
☞ c. L_1-<H_2> $\Delta F_0 (LH_1)$ = 20 Hz 222 ms-<107 ms>			*	ISD = 0.08

(Note that word-final tones in angle brackets are not analyzed.)

It can be seen from the above tableau that the reason why the rising contour is not faithfully rendered is arguably attributed to the proposed dispreference for the increasing degree of the steepness from citation (i.e. the inferred input) to sandhi. Candidates (a-b) are ruled out due to the very fact that slope is *not* decreasing in the sandhi output. So the rising contour of the initial tone is flattened to a low level tone, as in shown candidate (c). As a reminder, we have discussed in (45) that 20 Hz is the maximal tolerance range for level tones in LT.

So far, I have explained why the rising tone has to change to a level tone in sandhi. It turns out that the ban on smooth-to-steep mappings is the driving force of complete contour reduction for these combinations, in particular, the LH in the word-initial long syllables.

We are now in a position to deal with word-final rising tones, to which I turn in the subsequent section.

5.4.6 The Word-final Rising Tones

This section is concerned with the word-final rising tones. As we have seen, there are two conditions for the word-final rising tones: i) the initial tone is L-registered and ii) the final syllable is long. I will argue that *some* instances of the word-final rising tone are better treated as 'contour displacement' or 'contour movement.' Contour displacement refers to the phenomenon whereby a contour tone moves as a unit to another host syllable in the output (cf. Zhenhai Chinese in Li 2003). The rising tone in initial *short* syllables is displaced to the word-final *long* syllable. The driving force is contour preservation (RELCORR($x<y$)$_{\text{NUC}}$). The idea is that if there is a long syllable available, the rising tone moves to that syllable in order to satisfy this contour correspondence constraint. Consequently, we need to rank RELCORR($x<y$)$_{\text{NUC}}$ over MATCH-SLOPE$_{\text{ISD}\rightarrow 0}$. Note that this re-ranking does not undermine our previous analyses because these two constraints do not decide the winner in the preceding tabeleaux.

(56) Short^{LH}-Long^{T} → Short^{L}-Long^{LH} (where T=L or LH)

Citation: LH$_1$-<T$_2$> ΔF_0=30 Hz Rime=160 ms	MATCH-SLOPE$_{\text{ISD}>0}$	RELCORR ($x<y$)$_{\text{NUC}}$	MATCH-SLOPE$_{\text{ISD}\rightarrow 0}$	
☞ a. **<L->[18]-LH$_1$** ΔF_0(LH$_1$)=30Hz <108 ms>-238 ms			*	ISD = 0.17
b. LH$_1$-<T$_2$> ΔF_0=30 Hz 108 ms-<238ms>	*!		*	ISD = -0.17
c. L$_1$-<T$_2$> ΔF_0=20 Hz 108 ms-<238ms>		*!		ISD = 0

We see in candidate (b) that a smooth-to-steep slope mapping occurs. This output is again penalized by the slope correspondence constraint against these kinds of slope mapping. So, as discussed above, the initial LH undergoes contour reduction to avoid the dispreferred slope

---

[18] Note that in candidate (a), the boundary low tone L- is selected because of the ranking: *NON-HD/H » *NON-HD/L (see also the discussion in §5.4.4). Recall that the initial syllable is unstressed so we expect that a low level tone emerges in case of no input form.

mappings in candidate (c). But this time, the 'smaller than' relation ($x<y$) is lost in this candidate, hence is eliminated by the contour correspondence constraint.

We have mentioned that in these kinds of combination, i.e. short syllable-long syllable, the final long syllable offers a 'landing site' for the initial LH. Notice that in isolation, the (initial) LH is realized on a short syllable. So we expect a steep-to-smooth slope mapping if the initial LH is displaced to a long syllable. The prediction is borne out. In candidate (a), the ISD is 0.17, indicating that it's a steep-to-smooth mapping (i.e. ISD>0) and at the same time, contour preservation is satisfied.

Let us now consider the other context in which the final rising tone is attested: Long-Long combinations with two LH's in a row also surface as L-LH. As a reminder, the duration of a word-initial long syllable is in general 200 ms long. We have seen in (55) that this length is insufficient for a steep-to-smooth slope mapping. With regard to the word-final long syllable, however, its mean duration is considerably lengthened (=256 ms), in comparison to the word-initial long syllable (around 200 ms). One plausible account is that since L-registered long syllables are of the greatest duration in isolation (cf. Figure 5-9), it may not be surprising to see that this subtle durational difference is maintained in sandhi context. This comparatively longer duration in word-final position ensures that no smooth-to-steep mapping occurs in this combination.

(57) $\text{Long}^{LH}\text{-Long}^{LH} \rightarrow \text{Long}^{L}\text{-Long}^{LH}$

Citation:$\langle LH_1\rangle$-LH_2 ΔF_0 (LH_2) = 40 Hz Rime = 352 ms	MATCH-SLOPE$_{ISD>0}$	RELCORR $(x<y)_{NUC}$	MATCH-SLOPE$_{ISD\rightarrow 0}$	
☞ a. $\langle L_1\rangle$-LH_2 ΔF_0 (LH_2)=30 Hz (205 ms-)256 ms		*		ISD≈0
b. $\langle L_1\rangle$-H_2 ΔF_0 (LH_2)=20 Hz (205 ms-)256 ms		*!*	*	ISD=0.19

I have shown in (55) that word-initial long syllables cannot host a LH. So we can safely assume that LH invariably becomes L word-initially. In candidate (a), realizing a 30 Hz rise on a 256 ms rime does not violate any slope correspondence constraint. Instead, this output turns out to be a perfect match in slope. The contour preservation constraints are violated twice in candidate (b). So the word-final rising tone survives in this combination.

In sum, I have demonstrated that word-final rising tones occur in the following contexts: i) a final long syllable having a sufficient phonetic length or ii) contour displacement from a short syllable to a long syllable. Either way, the contour-licensing condition hinges on satisfaction of the slope correspondence constraint: MATCH-SLOPE$_{ISD>0}$.

5.4.7 The Absence of Word-final Rising Tones in Long Syllables

Lastly, the underlying rising tone in final syllables does not surface when the preceding syllable is H-registered. For the case at hand, the mean duration for the final syllable is 246 ms, which is slightly shorter than that of a rising-toned final syllable: 256 ms (cf. (57)). The ISD value is nevertheless negative if the rising tone is realized, as can be seen in candidate (a) below. Indeed, -0.04 is very tiny and may be ignored. But notice that the preceding long syllable is H-registered and we have learned that H-register syllables are invariably shorter than L-register syllables. We may speculate that in Long^{LH}-Long^{LH} combinations, both long syllables are L-registered so that phrase-final lengthening is strengthened, while for Long^{H}-Long^{LH} combinations, the duration of the final syllables are not awarded extra length because the preceding H-registered long syllable is not inherently longer. More research is in need to understand the exact mechanism for duration in LT. For the present purpose, however, it is important to note that all else being equal, a 238 ms-long rime *is* able to bear a rising tone (recall the contour displacement analysis in (56)). So the non-realization of the rising tone on a 246 ms-long rime cannot be attributed to a duration-based account, suggesting that slope correspondence is the major factor of contour tone licensing in LT.

(58) Long^{H}-Long^{LH} →Long^{H}-Long^{H}

Citation:<H$_1$>-LH$_2$ $\Delta F_0 = 40$ Hz Rime = 352 ms	MATCH-SLOPE$_{ISD>0}$	RELCORR $(x<y)_{NUC}$	MATCH-SLOPE$_{ISD\rightarrow 0}$	
a. <H$_1$>-LH$_2$ ΔF_0 (LH$_2$)= 30 Hz <174 ms>-246 ms	*!			ISD = -0.04
☞ b. <H$_1$>-H$_2$ ΔF_0(LH$_2$)= 20 Hz <174 ms>-246 ms		*	*	ISD = 0.13

This completes our analysis of the disyllabic tone sandhi in LT. The phenomenon in question can be characterized by the following statement: contour tone licensing is strictly constrained by slope correspondence, i.e. *no smooth-to-steep slope mappings*. This generalization gains support from the following asymmetries:

- The non-realization of a word-initial rising tone is not an artifact of tonal markedness because an average duration of 200 ms is sufficient for a rise of 30 Hz in initial sandhi position.

- The rising tone can appear on a 238 ms-long rime (i.e. contour displacement in (56)) but is banned from surfacing on a 246 ms-long rime (i.e. (58)). This is because the former is of the steep-to-smooth slope mapping (from short syllable to long syllable) and the latter is of the *smooth-to-steep* mapping (from isolation long syllable to (shortened) word-final long syllable).

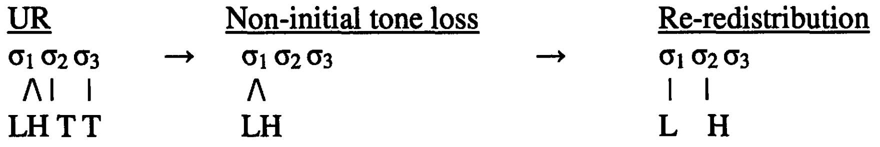
In summary, I have demonstrated in the foregoing sections that the proposed characteristics of slope correspondence, namely, i) slope identity preservation and ii) no increasing slope in output, must be factored into a decent account for tone sandhi of two remotely related languages, Hangzhou Chinese and Lhasa Tibetan. So far our discussion is restricted to the tone mapping within the syllable domain. In a wider context, I address the use of slope correspondence across syllables in the following section.

5.5 Bounded Tone Extension

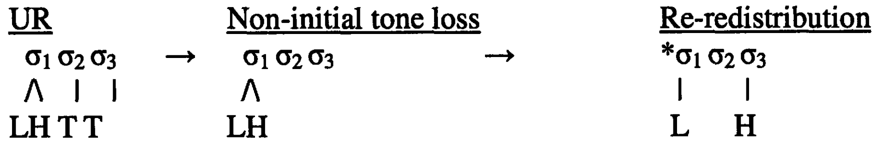
Unlike most segmental features, it is wellknown that tone may move several syllables away from its lexical source. In Chizigula (Kenstowicz and Kisseberth 1990), for instance, high tone migrates from the verb root to the penultimate syllable of the word, as the following example illustrates: lómbez 'to request' → ku-lombez-ez-án-a 'to request for each other.' (Note that low tone is unmarked). Of course, it is also well-known that low tone in Bantu languages is in general phonological inert, so that long-distance tone displacement might be made possible. In this section, I investigate the other side of the coin: "bounded tone extension." The phenomenon in question can be derivationally exemplified as follows.

(59) “Bounded” tone extension

a. Well-attested tone re-distribution



b. *Unattested* tone re-distribution



The observation is the following. If the non-initial tones fail to surface (here I use a theory-neutral term, ‘non-initial tone loss’), it is conceivable that the non-initial syllables may attract some tone from the initial syllable (again, I use a theory-neutral term,‘re-distribution’). Interestingly enough, the re-distribution pattern in (b), to the best of my knowledge, is unattested. The goal of this section is to show that slope correspondence offers a satisfactory account for the ‘bounded tone extension’ phenomenon.

5.5.1 Polysyllabic Tone Sandhi in Wu Chinese

The illustration of bounded tone extension in (59) is in fact the general schema of the tone sandhi in polysyllabic compounds in a handful of Wu Chinese dialects. The overall tone sandhi patterns can be sketched as follows. Non-initial tones are determined by the initial tones. The initial tone splits into two parts, superimposing itself on the entire tone sandhi domain. Let us now consider the polysyllabic tone sandhi patterns in one of the representative varieties of Wu Chinese, Shanghai Chinese. The (slightly revised) data are from Zee and Maddieson’s (1980) acoustic study.

(60) Polysyllabic tone patterns with respect to **INITIAL** tones in Shanghai Chinese

	Monosyllabic	Disyllabic	Trisyllabic	Quadrisyllabic
Tone 1	HL	H-L	H-M-L	H-M-Lꜛ-L
Tone 2	MꜛM	M-ꜛM	M-**H**-L	M-**H**-M-L
Tone 3	LM	L-ꜛM	L-**H**-L	L-**H**-M-L
Tone 4	$\underline{\text{H}}$	$\underline{\text{H}}$-H	$\underline{\text{H}}$ꜜ-H-L	$\underline{\text{H}}$ꜜ-H-M-L
Tone 5	$\underline{\text{LꜛM}}$	$\underline{\text{L}}$-LꜛM	$\underline{\text{L}}$-L-LMꜛ	$\underline{\text{L}}$-H-M-L

(Checked tones are underlined.)

In trisyllabic and quadrisyllabic compounds, all the contours end with L, suggesting that there is a low tonal target, or a phrasal low tone L- (Let us ignore the case of Tone 5, which features contour displacement to the final syllable in disyllablic and trisyllablic words). Of particular interest is that the H peaks of Tone 2 and Tone 3 (in boldface above) stay on the second syllable. Observe now the following illustrations of F_0 curves and duration for the trisyllabic tone pattern: LM-T-T → L-**H**-L (where T=any lexical tone). 12 tokens read by one male speaker in isolation were extracted from Ping's (2003) *The Phonetic Database of Shanghai Chinese*.

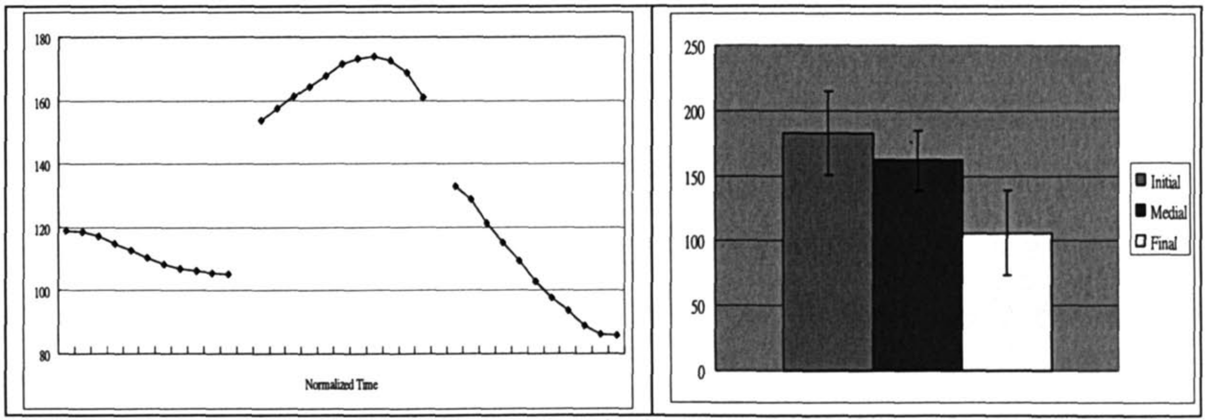


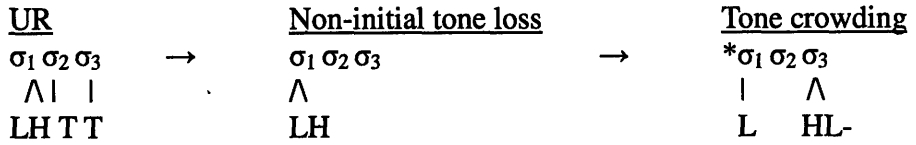
Figure 5-11 Normalized F_0 (in Hz) and duration (in ms) for the trisyllabic pattern LM-T-T → L-H-L in Shanghai Chinese

It can be seen from the above pitch tracing that the H peak (i.e. the F_0 maximum) is positioned in about 60% of normalized duration of the second syllable. The pitch contour in the final syllable is clearly a falling transition from the H peak in the second to the phrasal low tone L-. Therefore, we draw the conclusion that the underlying tone sequence /LM-T-T/ is mapped to L-H-L in the wake of complete tone neutralization on non-initial syllables. See Zee and Maddieson (1980) for a comprehensive illustration of polysyllabic tone patterns.

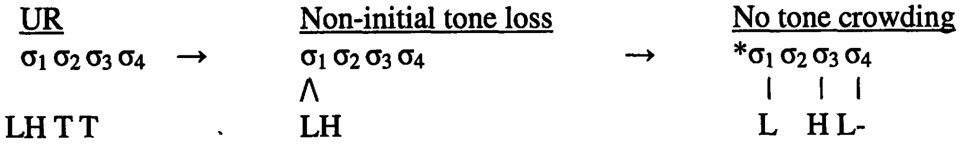
In trisyllabic compounds, it is possible that tone crowding in (61)a is the driving force for bounded tone extension. This is because the boundary low tone L- is realized on the final syllable. If H extends into the final syllable, forming a falling tone, there will be too many tones on the final syllable. Furthermore, recall from Figure 5-11 that the word-final syllable is normally of the shortest duration in tone sandhi domain. Taken together, these factors impede bounded tone extension.

(61) Bounded tone extension in Shanghai Chinese

a. Trisyllabic words



b. Quadrisyllabic words



However, tone crowding can be avoided in quadrisyallbic words. We see in (61)b that H migrates two syllables away from its lexical source (cf. (60)). Still, this move is not attested. Instead, H stays on the second syllable in surface realization.

In summary, tone re-distribution features 'local migration.' In comparison to long-distance migration in Chizigula, it seems that bounded tone extension constitutes a well-known property of tone: mobility. More importantly, at least both in Chizigula and Shanghia Chinese, high tone moves across toneless syllables (although their 'tonelessness' is motivated in very different way). Some attempts at explanation are made below. First, local migration can be satisfaction of tonal alignment constraint, as illustrated in the following tableau.

(62) An alignment account

$\sigma_1 \sigma_2 \sigma_3 \sigma_4$ Λ| | | LHT TT	LINEARITY	ALIGN-L(H, PRWD)
☞ a. $\sigma_1 \sigma_2 \sigma_3 \sigma_4$ | | LH L-		*
b. $\sigma_1 \sigma_2 \sigma_3 \sigma_4$ | | L H L-		*!*
c. $\sigma_1 \sigma_2 \sigma_3 \sigma_4$ | | H L L-	*!	

This ranking of constraints predicts the desired result. We see in candidate (b) that if H docks on the third syllable, ALIGN-L(H, PRWD) is violated twice (Let us simply assume the gradient version of alignment constraint. See McCarthy (2003b) for more discussion). Candidate (c) satisfies ALIGN-L(H, PRWD) but is penalized by LINEARITY, the anti-metathesis constraint. Employing alignment is problematic in that unattested patterns are predicted: if we flip the current ranking, candidate (c) will be selected as the winner. To the best of my knowledge, tone re-distributions of this sort are unattested.

Invoking contiguity is hopeless, too. A contiguity-based account is provided in the following tableau. Both candidates (a-b) satisfy 'No skipping,' i.e. INPUT-CONTIGUITY-(TONE) (abbr. I-CONTIG-(T)) because the output tone sequence is also a contiguous string in the input (modulo the boundary low tone L-).

(63) A contiguity account

$\sigma_1 \sigma_2 \sigma_3 \sigma_4$ Λ | || LH T T T	I-Contig-(Tone)
☞ a. $\sigma_1 \sigma_2 \sigma_3 \sigma_4$ | | L H L-	
☞ b. $\sigma_1 \sigma_2 \sigma_3 \sigma_4$ | | L H L-	

Li (2003: 70) proposes that "the tonal target is realized either in its host syllable or in the immediately following syllable." This idea is formalized as the following constraint:

(64) Contiguity(Syl-Tone)

> 'No syllable intervenes between the syllable that bears the tone in the input and the syllable that realizes the tone in the output.'

This constraint predicts the desired result because we see, for example, in candidate (63)b that there is an intervening syllable between L and H. But the logic of this constraint is questionable. Notice that contiguity is also defined in terms of individual element-based correspondence theory. The formulation of Contiguity(Syl-Tone) treats 'syllable and its associated tone' as an entity. So if there is a toneless syllable intervening between two tone-bearing syllables, for example, candidate (63)b, this *toneless* syllable (i.e. σ_2 in (63)b) should not count as an entity that is subject to evaluation of Contiguity.

I argue that slope correspondence provides a more straightforward account for the 'bounded tone extension' phenomenon. As we can see in the following tableau, candidate (b) will be always harmonically bound, provided that MATCH-SLOPE is active.

(65) A slope correspondence account

<table>
<tr><th>$\sigma_1 \sigma_2 \sigma_3 \sigma_4$
Λ | | |
LHT TT</th><th>MATCH-SLOPE$_{\text{ISD}\rightarrow 0}$</th></tr>
<tr><td>☞ a. $\sigma_1 \sigma_2 \sigma_3 \sigma_4$
| |
LH L-</td><td></td></tr>
<tr><td>b. $\sigma_1 \sigma_2 \sigma_3 \sigma_4$
| |
L H L-</td><td>*!</td></tr>
</table>

In summary, bounded tone extension is robustly attested in a handful of Wu Chinese languages. I have shown that rightward local migrations of this sort should be better treated as slope correspondence, rather than satisfaction of alignment or contiguity. In the following section, let us look at yet another parallel instantiation in Mandarin Chinese.

5.5.2 Neutral tone in Mandarin Chinese

In Mandarin Chinese, when there are three neutral tones (toneless syllables) in a row, preceded by a syllable specified with a lexical tone, we find a parallel phenomenon with the Wu Chinese data just discussed in the preceding section. Observe now the following F_0 tracings, where Tone 1 = H, Tone 2 = R, Tone 3 = L, Tone 4 = F, N1/2/3 = the first/second/third neutral tone. The data are taken from Li's (2003) experimental results.

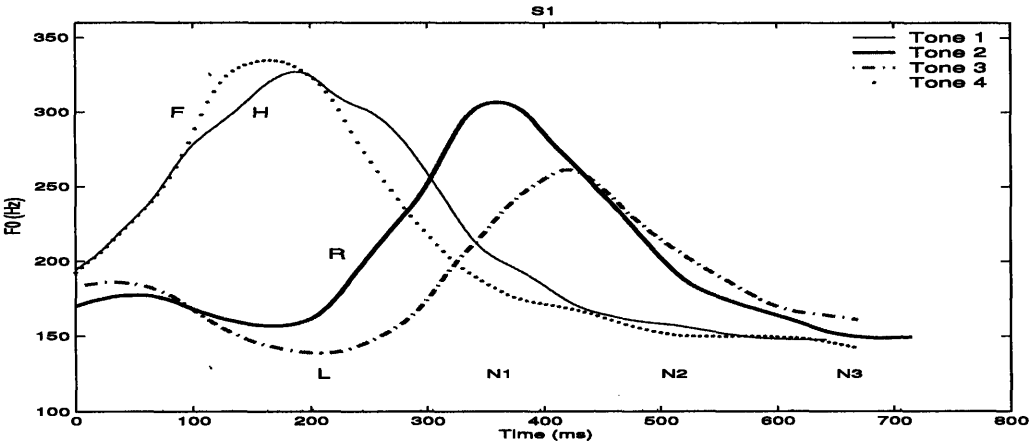
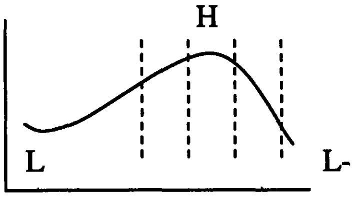


Figure 5-12 Neutral tone sequences in Mandarin (Adapted from Li 2003)

As seen, all F_0 curves converge towards the endpoint of N_3, suggesting that there is a low tonal target, or a boundary low tone L-. Of particular interest is the fact that the H peaks of Tone 2 (R) and Tone 3 (L) as well as the low target of Tone 4 (F) do not extend to N2. Instead, tone extension "stops" in N1. Again, bounded tone extension is attested in Mandarin Chinese.

(66) Why excessive high peak delay is *not* allowed?



(dotted line=syllable boundary)

Suppose that H peak docks on N2 and the low boundary tone L- is realized on N3. It is reasonable to say that tone crowding is not at issue. In addition to this, the arguments against the alignment, the contiguity and Li's (2003) CONTIGUITY(SYL-TONE) are all extendable to the Mandarin Chinese data without any problem. Consequently, we arrive at a similar conclusion: excessive high peak delay is banned because this output form is always harmonically bound given that MATCH-SLOPE is active. The analysis is illustrated in the following tableau.

(67) A slope correspondence account

<table>
<tr><td>$T_1 N_2 N_3 N_4$
Λ
LH</td><td>MATCH-SLOPE$_{ISD\rightarrow 0}$</td></tr>
<tr><td>☞ a. $T_1 N_2 N_3 N_4$
| |
L H L-</td><td></td></tr>
<tr><td>b. $T_1 N_2 N_3 N_4$
| |
L H L-</td><td>a>b</td></tr>
</table>

(where T=lexical tone, N=neutral tone, where a>b means b is harmonically bound by a.)

5.5.3 Summary of this Section

In this section, I provided two case studies of the 'bounded tone extension' phenomenon. I have demonstrated that bounded tone extension is best analyzed in terms of slope correspondence. At the same time, the current approach also predicts that long-distance tone migration is possible as long as slope-matching is not necessary. For example, non-local tone displacement or tone attraction in West African languages normally involves high level tone only. I know of no documented case where falling or rising tones in West African languages behave this way.

In summary, this typological disparity of tone languages, i.e. long-distance tone displacement vs. bounded tone extension, may boil down to the presence of absence of the slope-matching mechanism.

5.6 Concluding Remarks

Slope correspondence is not a brand new idea in the tone sandhi literature. Researchers have already pursued similar ideas in analytical practice. For example, Yip (2002) uses 'PRES(ERVE)-(LH)' to derive the Lhasa Tibetan disyllabic tone sandhi phenomenon whereby the underlying LH tonal melody of Syllable S is realized *as a whole* on the same syllable in the output. That is, LH should not split and re-associate to two distinct syllables on surface (cf. §5.4). In his analysis of Shaoxing Chinese (see also chapter 4), Zhang (2006) proposes 'IDENT-R(ISE)' to ensure that the tonal specification of an underlying rising tone should be identical in the input and output. In view of unit correspondence (McCarthy and Prince 1995), these diagnostics are an unwarranted stipulation since the contour between two tones, be it rising or falling, is not defined in the original formulation. More precisely, individual, element-based correspondence is unable to adequately express 'relational identity' of two successive tones.

On the other hand, slope correspondence has also been treated as satisfaction of tonal alignment constraints. In his discussion of the bitonal accent H*+L (Accent 2; here I use the asterisk to represent the underlyingly associated H tone) in Stockholm Swedish, Gussenhoven (2004) interprets Riad's (1998) CONCATENATE as 'tones in bitonal morphemes are aligned with each other' and accordingly gives an explanation of the tone pattern whereby the *unassociated* trailing L of Accent 2 is typically realized immediately after H*. Under the pressure from CONCATENATE, L's left edge is aligned with the right edge of H*. This said, one lurking problem is that all else being equal, it is predicted that in some bitonal accent language, L's right edge may be aligned with the left edge of H*, yielding a metathesized Accent 2: L+H*. To the best of my knowledge, this hypothetical case is unattested. Instead, the effect of CONCATENATE is better treated as faithfulness, rather than markedness. It is conceivable that slope correspondence fits in the characterization given above: slope identity of the correspondent H*+L sequences (between the inferred input and surface form) should be as similar as possible so that the unassociated L occurs immediately after H*.

In this chapter, I have formulated the slope correspondence constraints and provided a unified account for an array of data from various languages. In addition to duration and the neighboring tones, slope correspondence should be regarded as another contour tone licensing condition. This claim is evidenced in the general dispreference for smooth-to-steep slope mappings even though mappings of this sort are articulatorily feasible and are more faithful to the input pitch excursion size. Furthermore, slope correspondence is not restricted to the syllabic nucleus domain. The 'bounded tone extension' phenomenon in East Asian languages indicates that rightward local migration of the rightmost portion of a contour tone is strictly constrained by locality. This locality effect is difficult or problematic to analyze with the standard OT approach such as alignment or contiguity. As we have discussed, slope correspondence plays a key role in this regard and the typological difference between the African-type and Asian-type tone displacement is now attributable to the presence and absence of the slope-matching requirement.

References

Abramson, Arthur. 1978. Static and dynamic acoustic cues in distinctive tones. *Language and speech* 21:319-325.

Akanlig-Pare, George, and Michael Kenstowicz. 2002. Tone in Buli. *Studies in African Linguistics* 31:55-95.

Archangeli, Diana, and Douglas Pulleyblank. 1994. *Grounded Phonology*. Cambridge, Mass.: MIT Press.

Bao, Zhiming. 1990. On the Nature of Tone, MIT: Ph. D. Dissertation.

Bao, Zhiming. 1999. *The Structure of Tone*. New York: Oxford University Press.

Benua, Laura. 1997. Transderivational Identity: Phonological Relations between Words, University of Massachusetts, Amherst: Ph. D. Dissertation.

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

Boersma, Paul, and David Weenink. 2007. Praat: doing phonetics by computer (Version 4.6.09) [Computer program]. Retrieved June 24, 2007 from http://www.praat.org/.

Broselow, Ellen. 1992. The structure of fricative-stop onsets. Ms. State University of New York, Stony Brook: Stony Brook, NY.

Cahill, Michael. 2007. More universals of tone. *SIL Electronic Working Papers 2007-007*:1-10.

Cao, J., and I. Maddieson. 1992. An exploration of phonation types in Wu dialects of Chinese. *Journal of Phonetics* 20:77-92.

Causley, Trisha. 1997. Identity and featural correspondence: The Athapaskan case. In *Proceedings of the North East Linguistic Society 27*, ed. Kiyomi Kusumoto, 93-105. Amherst, Mass.: GLSA Publications.

Chao, Y.-R. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chao, Yuan-ren. 1930. A system of tone letters. *Le Maître Phonétique* 45:24-27.

Chen, M. Y. 2000. *Tone Sandhi: Patterns Across Chinese Dialects*: Cambridge University Press.

Chen, Y., and Y. Xu. 2006. Production of Weak Elements in Speech: Evidence from F0 Patterns of Neutral Tone in Standard Chinese. *Phonetica* 63:47-75.

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.

Clark, Mary. 1978. A Dynamic Treatment of Tone with Special Attention to the Tonal System of Igbo, University of New Hampshire: Ph. D. Dissertation [Distributed by Indiana

University Linguistics Club].

Clements, G. N. 1979. The description of terraced-level tone languages. *Language* 55:536-558.

Clements, G. N., and K. C. Ford. 1979. Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry* 10:179-210.

Clements, G. N., and Samuel Jay Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable*. Cambridge, Mass.: MIT Press.

Connell, B. 2000. The Perception of Lexical Tone in Mambila. *Language and Speech* 43:163-182.

de Lacy, Paul. 2002a. The Formal Expression of Markedness, University of Massachusetts, Amherst: Ph. D. Dissertation.

de Lacy, Paul. 2002b. The interaction of tone and stress in Optimality Theory. *Phonology* 19:1-32.

De Lacy, Paul. 2007. The formal properties of phonological precedence. Paper presented at *Conference on precedence relations*, City University of New York.

Dilley, Laura C. 2004. The Phonetics and Phonology of Tonal Systems, MIT: Ph. D. Dissertation.

Duanmu, S. 1992. An autosegmental analysis of tone in four Tibetan languages. *Linguistics of the Tibeto-Burman Area* 15:65-91.

Duanmu, S. 1993. Rime length, stress, and association domains. *Journal of East Asian Linguistics* 2:1-44.

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

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

Duanmu, San. 2000. *The Phonology of Standard Mandarin*. Oxford: Oxford University Press.

Dyk, Anne, and Betty Stoudt eds. 1965. *Vocalbulario Mixtec de San Miguel el Grande*. D.F., Mexico: El Instituto Linguistico de Verano.

Elimelech, Baruch. 1978. *A Tonal Grammar of Etsako*: UCLA Working Papers in Phonetics 35. Los Angeles, CA: University of California, Los Angeles.

Esling, John. 1994. University of Victoria Phonetic Database (Version 3). Victoria, BC: Speech Technology Research, Ltd..

Fleischhacker, Heidi. 2001. Cluster-dependent epenthesis asymmetries. In *UCLA Working Papers in Linguistics (Papers in Phonology 5)*, eds. Adam Albright and Taehong Cho, 71-116. Los Angeles, CA: Department of Linguistics, University of California, Los Angeles.

Fleischhacker, Heidi. 2005. Similarity in Phonology: Evidence from Reduplication and Loan Adaptation, University of California, Los Angeles: Doctoral dissertation.

Flemming, Edward. 1995. Auditory Representations in Phonology, UCLA: Ph. D. Dissertation.

Flemming, Edward. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18:7-44.

Flemming, Edward. 2002. *Auditory representations in phonology*. New York: Routledge.

Flemming, Edward. 2006. The role of distinctiveness constraints in phonology. Ms. MIT, Cambridge, MA.

Fok-Chan, Yuen-Yuen. 1974. *A perceptual study of tones in Cantonese*. Hong Kong: Centre of Asian Studies, University of Hong Kong.

Fu, Qian-Jie, and Fan-Gang Zeng. 2000. Identification of temporal envelope cues in Chinese tone recognition. *Asia Pacific Journal of Speech, Language and Hearing* 5:45-57.

Gandour, Jackson, and Richard Harshman. 1978. Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and Speech* 21:1-33.

Gandour, John T. 1978. The perception of tone. In *Tone: A linguistic survey*, ed. Victoria Fromkin, 41-76. New York: Academic Press.

Gerfen, Chip. 1999. *Phonology and Phonetics in Coatzospan Mixtec*. Dordrecht: Kluwer Academic Publishers

Goldsmith, John. 1976. Autosegmental Phonology, MIT: Ph.D. Dissertation.

Goldsmith, John. 1990. *Autosegmental and Metrical Phonology*. Oxford and Cambridge, Mass.: Blackwell.

Gordon, M. 2001. A typology of contour tone restrictions. *Studies in Language* 25:423-462.

Gordon, M., and P. Ladefoged. 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29:383-406.

Gouskova, Maria. 2004. Relational hierarchies in Optimality Theory: the case of syllable contact. *Phonology* 21:201-250.

Gussenhoven, C. 2004. *The Phonology of Tone and Intonation*: Cambridge University Press.

Hallé, Pierre. A. 1994. Evidence for tone-specific activity of the sternohyoid muscle in Modern standard Chinese. *Language and speech* 37:103-123.

Hansson, Gunnar Olafur. 2001. Theoretical and Typological Issues in Consonant Harmony, University of California, Berkeley.

Haraguchi, Shosuke. 1977. *The Tone Pattern of Japanese: An Autosegmental Theory of Tonology*. Tokyo: Kaitakusha.

Heinz, Jeffrey. 2005. Reconsidering Linearity: Evidence from CV Metathesis. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, eds. John Alderete et al., 200-208. Somerville, MA: Cascadilla Proceedings Project.

Hoffmann, Carl. 1963. *A Grammar of the Margi Language*. London: Oxrford University Press.

Hollenbach, Barbara. 1977. Phonetic vs. phonemic correspondence in two Trique dialects. In *Studies in Otomanguean Phonology*, ed. W. Merrifield, 35-68. Dallas: Summer Institute of Linguistics.

Hollenbach, Barbara. 1984. The phonology and morphology of tone and laryngeals in Copala Trique, University of Arizona: Ph. D. Dissertation.

Hombert, J. M. 1978. Consonant types, vowel quality, and tone. *Tone: A Linguistic Survey*:77-112.

Hombert, J. M., J. J. Ohala, and W. G. Ewan. 1979. Phonetic Explanations for the Development of Tones. *Language* 55:37-58.

House, David. 1996. Differential perception of tonal contours through the syllable. *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on* 4.

Hume, E., and G. Tserdanelis. 2003. Labial unmarkedness in Sri Lankan Portuguese Creole. *Phonology* 19:441-458.

Hyman, L. M., and K. VanBik. 2004. Directional Rule Application and Output Problems in Hakha Lai Tone. *Special Issue: Phonetics and Phonology, Language and Linguistics* 5:821-861.

Hyman, Larry, and Russell Schuh. 1974. Universals of tone rules: Evidence from West Africa. *Linguistic Inquiry* 5:81-115.

Hyman, Larry. 1993. Register tones and tonal geometry. In *The Phonology of Tone: The Representation of Tonal Register*, eds. K. Snider and Harry van der Hulst, 75-108. Berlin: Mouton de Gruyter.

Hyman, Larry. to appear. Universal of tone rules: 30 years later. In *Tones and Tunes: Studies in Word and Sentence Prosody*, ed. Carlos Gussenhoven and Tomas Riad. Berlin: Mouton de Gruyter.

Iwata, R., H. Hirose, S. Niimi, and S. Horiguchi. 1991. Physiological properties of" Breathy" phonation in a Chinese dialect-A Fiberoptic and Electromyographic study on Suzhou dialect. *Proceedings of the XIIth International Congress of Phonetic Sciences* 3:162.

Jun, Jongho. 2002. Positional faithfulness, sympathy, and inferred input. Ms., *Kyongsan, South Korea: Yeungnam University*.

Kaji, Shigeki. 1996. Tone reversal in Tembo (Bantu J.57). *Journal of African Languages and Linguistics* 17:1-26.

Kang, Yoonjung. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20:219-273.

Kenstowicz, M., and C. Kisseberth. 1990. Chizigula Tonology: the Word and Beyond, ed. Sharon Inkelas and Draga Zec (1990), 163-194.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.

Kenstowicz, Michael. 2006. Tone loans: the adaptation of English loanwords into Yoruba. In *Selected Proceedings of the 35th Annual Conference on African Linguistics*, eds. John

Mugane et al., 136-146. Somerville, MA: Cascadilla Proceedings Project.

Kiparsky, Paul. 1973. How abstract is phonology? In *Three Dimensions of Linguistic Theory*, ed. O. Fujimura, 5-56. Tokyo: TEC.

Klatt, D. H. 1973. Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception. *The Journal of the Acoustical Society of America* 53:8.

Kong, Jiangping. 1992. Ziyun Miaoyu wupingdiao xitong de shengxue ji ganzhi yanjiu [Acoustic and perceptual studies of the five level tones in Ziyun Miao] In *Minzu Yuwen Yanjiu Xintan [New Explorations on Minority Languages]*, ed. Society of Chinese Minority Languages, 152-163. Chengdu: Sichuan Nationalities Press.

Kuo, Y.-C., Yi Xu, and Moira Yip. in press. The phonetics and phonology of apparent cases of iterative tonal change in Standard Chinese. Paper presented at *Proceedings of TIE Conference*.

Lamontagne, Greg, and Keren Rice. 1995. A correspondence account of coalescence. In *University of Massachusetts Occasional Papers in Linguistics 18*, eds. Jill Beckman et al., 211-224. Amherst, Mass.: GLSA Publications.

Lau, Sze-lok. 2002. Tone and Tone Sandhi in Suzhou, City University of Hong Kong: MPhil Thesis.

Leben, Will. 1973. Suprasegmental Phonology, MIT: Ph. D. Dissertation.

Li, Zhiqiang. 2003. The Phonetics and Phonology of Tone Mapping in a Constraint-based Approach, MIT.

Lombardi, L. 1998. Evidence for MaxFeature constraints from Japanese. *University of Maryland Working Papers in Linguistics* 7:41-62.

Lombardi, Linda. 1995. Why place and voice are different: Constraint-specific alternations and Optimality Theory. Ms. University of Maryland at College Park.

Lombardi, Linda. 2001. Why Place and Voice are different: Constraint-specific alternations in Optimality Theory. In *Segmental Phonology in Optimality Theory: Constraints and Representations*, ed. Linda Lombardi, 13-45. Cambridge: Cambridge University Press.

Lovins, Julie. 1971. Melodic conspiracies in Lomongo tonology. In *Papers from the 7th Regional Meeting, Chicago Linguistic Society*, eds. Douglas Adams et al., 469-478. Chicago: Chicago Linguistic Society.

Macaulay, M. A. 1996. *A Grammar of Chalcatongo Mixtec*: University of California Press.

Maddieson, Ian. 1978. Universals of tone. In *Universals of Human Language, vol. 2: Phonology*, ed. Joseph Greenberg, 335-365. Stanford: Stanford University Press.

McCarthy, J. 2003a. Comparative markedness. *Theoretical Linguistics* 29:1-51.

McCarthy, John J. 1986. OCP Effects: Gemination and antigemination. *Linguistic Inquiry*

17:207-263.

McCarthy, John J., and Alan Prince. 1993. *Prosodic morphology I: Constraint interaction and satisfaction*: Rutgers Technical Report TR-3. New Brunswick: Rutgers University Center for Cognitive Science.

McCarthy, John J., and Alan Prince. 1994. The emergence of the unmarked: Optimality in prosodic morphology. In *Proceedings of the North East Linguistic Society 24*, ed. Mercè Gonzàlez, 333-379. Amherst, Mass.: GLSA Publications.

McCarthy, John J., and Alan Prince. 1995. Faithfulness and Reduplicative Identity. In *University of Massachusetts Occasional Papers in Linguistics 18*, eds. Jill Beckman et al., 249-384. Amherst, Mass.: GLSA Publications.

McCarthy, John J. 1999. Sympathy and phonological opacity. *Phonology* 16:331-399.

McCarthy, John J. 2002. Comparative Markedness. Ms. University of Massachusetts, Amherst.

McCarthy, John J. 2003b. OT constraints are categorical. *Phonology* 20:75-138.

Minkova, Donka. 2003. *Alliteration and Sound Change in Early English*. Cambridge: Cambridge University Press.

Myers, James, and Jane Tsay. 2003. A formal functional model of tone. *Language and Linguistics* 4.1:105-138.

Odden, David. 1986. On the Obligatory Contour Principle. *Language* 62:353-383.

Odden, David. 1994. Adjacency parameters in phonology. *Language* 70:289-330.

Odden, David. 1995. Tone: African languages. In *The Handbook of Phonological Theory*, ed. John Goldsmith, 444-475. Cambridge, MA, and Oxford, UK: Blackwell.

Ohala, J. J., and W. G. Ewan. 1973. Speed of Pitch Change. *The Journal of the Acoustical Society of America* 53:345.

Ohala, John. 1978. The production of tone. In *Tone: A Linguistic Survey*, ed. Victoria Fromkin, 5-39. New York: Academic Press.

Ojo, Valentine. 1977. English-Yoruba Language Contact in Nigeria, Eberhard-Karls-Universitaet: Ph. D. Dissertation.

Ola Orie, Olanike, and Douglas Pulleyblank. 2002. Yoruba Vowel Elision: Minimality Effects. *Natural Language & Linguistic Theory* 20:101-156.

Parker, Steve. 2002. Quantifying the Sonority Hierarchy, University of Massachusetts, Amherst: Ph. D. Dissertation.

Pike, Kenneth. 1948. *Tone Languages*. Ann Arbor: University of Michigan Press.

Ping, Yueling. 2001a. Shaoxing fangyan shengdiao shiyan yanjiu [Experimental studies of Shaoxing Chinese tonal system]. In *Wuyu Shengdiao de Shiyan Yanjiu [Experimental studies of Wu Chinese Tonal Systems]*, ed. Rujie You and Jianqiao Yang, 97-130. Shanghai: Fudan University Press.

Ping, Yueling. 2001b. Shanghai shiqu fangyan shengdiao shiyan yanjiu [Experimental studies of Shanghai Chinese tonal system]. In *Wuyu Shengdiao de Shiyan Yanjiu [Experimental studies of Wu Chinese Tonal Systems]*, ed. Jianqiao Rujie You and Yang, 17-38. Shanghai: Shanghai Education Publishing House.

Prince, Alan, and Paul Smolensky. 1991. Connectionism and Harmony Theory in linguistics: Department of Computer Science, University of Colorado, Boulder.

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Technical Report 2: Center for Cognitive Science, Rutgers University.

Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, Mass., and Oxford, UK: Blackwell.

Pulleyblank, D. 1996. Neutral vowels in optimality theory: A comparison of yoruba and wolof. *The Canadian journal of linguistics* 41:295-347.

Pulleyblank, D. 2005. A note on tonal markedness in Yoruba. *Phonology* 21:409-425.

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

Qiao, Quansheng, and Li Chen. 2004. Pingyaohua Yindang [The Phonetic database of Pingyao Chinese]. In *Xiandai Hanyu Fangyan Yinku [The Phonetic Database of Modern Chinese Dialects]*, ed. Jingyi Hou. Shanghai: Shanghai Education Publishing House.

Ren, Nianqi. 1992. Phonation Types and Stop Consonant Distinctions: Shanghai Chinese, University of Connecticut: Ph. D. Dissertation.

Riggle, J. 1999. Relational Markedness in Bantu Vowel Height Harmony. Ms., *UCLA*. Los Angeles, CA.

Rose, Sharon, and Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80:475-532.

Schuh, Russell. 1978. Tone rules. In *Tone: A Linguistic Survey*, ed. Victoria Fromkin, 221-256. New York: Academic Press.

Shen, Tong. 1995. The underlying representation of Suzhou Tones. *New Asia Academic Bulletin* 11:129-143.

Shi, Feng , Lin Shi, and Rongrong Liao. 1987. An experimental analysis of the five level tones of the Gaoba Dong language. *Journal of Chinese Linguistics* 15:335-361.

Silverman, D. 2002. Laryngeal complexity in Otomanguean vowels. *Phonology* 14:235-261.

Silverman, Daniel. 1997. Tone Sandhi in Comaltepec Chinantec. *Language* 73:473-492.

Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Paper presented at *Rutgers Optimality Workshop -1*, New Brunswick, NJ.

Snider, Keith. 1990. Tonal upstep in Krachi: Evidence for a register tier. *Language* 66:453-474.

Snider, Keith. 1999. *The geometry and features of tone*: Summer Institute of Linguistics.

Steriade, Donca. 1988a. Reduplication and syllable transfer in Sanskrit and elsewhere.

Phonology 5:73-155.

Steriade, Donca. 1988b. Greek accent: a case for preserving structure. *Linguistic Inquiry* 19:271-314.

Steriade, Donca. 1997. Phonetics in Phonology. Paper presented at *SICOL 1997*, Seoul, Korea.

Steriade, Donca. 2001. Directional asymmetries in place assimilation: a perceptual account. In *The Role of Speech Perception in Phonology*, eds. Elizabeth Hume and Keith Johnson. San Diego: Academic Press.

Steriade, Donca. 2006. Contour correspondence: evidence from cluster interruption. Paper presented at *Old World Conference in Phonology 3*, Budapest, Hungary.

Sun, J. T. S. 1997. The Typology of Tone in Tibetan. *Chinese Languages and Linguistics IV: Typological studies of languages in China, Taibei, Taiwan, Symposium Series of the Institute of History and Philology-Academia Sinica*:485-521.

Sun, Ruixin. 2001. Suzhou Fangyan Biandia Shiyan Yanjiu [Experimental studies of Suzhou Chinese tone sandhi]. In *Wuyu Shengdiaode Shiyan Yanjiu [Experimental Studies of Wu Chinese Tone Sandhi]*, eds. Rujie You and Jianqiao Yang. Shanghai: Fudan University Press.

Sundberg, Johan. 1973. Data on maximum speed of pitch changes. *Speech Transmission Laboratory Quarterly Progress and Status Report* 4:39-47.

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

Suzuki, Keiichiro. 1998. A Typological Investigation of Dissimilation, University of Arizona: Ph. D. Dissertation.

Tranel, B. 1992/1994. Tone sandhi and vowel deletion in Margi. *Studies in African Linguistics* 23:111-183.

Tranel, Bernard. 1996. Rules vs. constraints: a case study. In *Current trends in Phonology: Models and Methods*, eds. J. Durand and B. Laks, 711-730. Paris: CNRS and Salford: University of Salford.

Tsay, Su-chuan Jane. 1994. Phonological pitch, University of Arizona: Ph.D. Dissertation.

van Oostendorp, Marc. 2005. The Theory of Faithfulness. Ms. Amsterdam, Meertens Instituut.

Walker, R. 2000. *Nasalization, neutral segments, and opacity effects*: Garland.

Wang, William S.-Y. 1967. Phonological features of tone. *International Journal of American Linguistics* 33:93-105.

Whalen, Douglas , and Yi Xu. 1992. Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica* 49:25-47.

Wolf, Matthew, and John McCarthy. 2007. Less than zero: Correspondence and the null output. In *Modeling Ungrammaticality in Optimality Theory*, ed. Curt Rice. London: Equinox

Publishing.

Xu, Y. 1994. Production and perception of coarticulated tones. *The Journal of the Acoustical Society of America* 95:2240.

Xu, Y. 1995. Production and perception of coarticulation tones. *J. Acoust. Soc. Am* 95.

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

Xu, Y. 1998. Consistency of Tone-Syllable Alignment across Different Syllable Structures and Speaking Rates. *Phonetica* 55:179-203.

Xu, Y., and Q. E. Wang. 2001. Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication* 33:319-337.

Xu, Y., and X. Sun. 2002. Maximum speed of pitch change and how it may relate to speech. *The Journal of the Acoustical Society of America* 111:1399.

Yip, M. 1993. Tonal register in East Asian languages. In *The Phonology of Tone: The Representation of Tonal Register*, 245-268.

Yip, Moira. 1980. The Tonal Phonology of Chinese, MIT: Ph. D. Dissertation.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Yip, Mora. 2001. Tonal features, tonal inventories and phonetic targets. In *UCL Working Papers in Linguistics 13*, eds. Corinne Iten and Ad Neeleman, 303-329. London: Department of Phonetics and Linguisitcs, University College London.

Zee, Eric, and Ian Maddieson. 1980. Tones and tone sandhi in Shanghai: Phonetic evidence and phonological analysis. *Glossa* 14:45-88.

Zhang, Jie. 1999. Duration in the tonal phonology of Pingyao Chinese. In *UCLA Working Papers in Linguistics 2 (Papers in Phonology 3)*, ed. Gordon Matthew, 147-206. Los Angeles, CA: Department of Linguistics, University of California, Los Angeles.

Zhang, Jie. 2002a. Non-contrastive features and categorical patterning in Chinese diminutive suffixation: MAX [F] or IDENT [F]? *Phonology* 17:427-478.

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

Zhang, Jisheng. 2006. *The phonology of Shaoxing Chinese*. Utrecht: LOT.

Zhou, Jiwen ed. 1983. *Zangwen Pinyin Jiaocai (Lasa Yin)/bod yig gi sgra sbyor slob deb (lha sa'i skad) [A Pronunciation Guide for Lhasa Tibetan]*. Beijing: The Nationalities Press.

Zhu, Xiaonong Sean. 1999. *Shanghai Tonetics*. Munich: Lincom Europa.

Zhu, Xiaonong Sean. 2006. Lun Fengyu Sidu Biaodiaozhi. In *Yinyu Yanjiu*, 170-187. Beijing: Shangwu Yishuguan.

Zuraw, Kie. 2005. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog. Ms., *UCLA*. Los Angeles, CA.

Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language* 83:277-316.