TONE SANDHI

Patterns across Chinese dialects

MATTHEW Y. CHEN

City University of Hong Kong

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011–4211, USA www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain

First published 2000

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times 10/13pt [GC]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data

Chen, Matthew Y.
Tone sandhi: patterns across Chinese dialects / Matthew Y. Chen.
p. cm. – (Cambridge studies in linguistics)
Includes bibliographical references.
ISBN 0 521 65272 3
1. Chinese language – Tone. I. Title. II. Series.
PL1213.C445 2000
495.1′16–dc21 99–44921 CIP

ISBN 0 521 652723 hardback

Contents

1 *Setting the stage*

This introductory chapter is intended to provide the necessary background for our investigation of tone sandhi. After a brief description of the genetic grouping of the languages of China, from which we draw the bulk of our primary data (section 1), I give a thumbnail sketch of the tone system of Middle Chinese (circa AD 600) and its evolution into the diverse patterns we see in modern dialects (sections 2–3). Historical tonal categories furnish us with a common frame of reference as we move from one dialect to another. I then set tone sandhi in the context of various types of tonal perturbations in connected speech, including tonal coarticulation, intonational effects, and morphologically conditioned tone changes (section 4). Tone sandhi processes often strike the analyst as arbitrary and totally lacking in phonetic or functional motivation. Section 5 shows that we can make sense of, if not explain, certain puzzling synchronic facts if we look at them from a diachronic perspective. This chapter closes with some terminological clarification (section 6).

1 Languages and dialects of China

According to *Major Statistics of the 1982 Census*, published by the People's Republic of China State Statistics Bureau (Beijing, October 1982), China (including Taiwan) has a population of 1,026 million.[1] Of these, 977.2 million or 95.2% speak one form or another of Chinese. The remaining 46.2 million are distributed over a wide variety of language families/stocks, spoken mostly on the periphery of China, with a high concentration of speakers of "minority" languages across the southwestern provinces. *Language Atlas of China* (Longman, Hong Kong 1987), compiled by the Australian Academy of the Humanities and the Chinese Academy of Social

[1] Quoting official statistics, *Language Atlas of China* (A-1) puts the population at 1.1 billion by April 1989.

Sciences, affords us a glimpse of the linguistic diversity within the political boundaries of China, that includes Sinitic as well as Tibeto-Burman, Kam-Tai, Miao-Yao, Austronesian (Formosan), Mon-Khmer, Altaic, and even Indo-European languages.

Of more immediate interest is the classification of the Chinese languages, more commonly referred to as "dialects." We can make meaningful typological generalizations not only about individual dialects, but about dialect groups. For instance, the "southern" dialects typically have larger tonal inventories than the Mandarin group (see Cheng 1973b, 1991 for statistical data). More importantly, the Jin, Wu, Min, Hakka, and some Mandarin dialects display highly complex tone sandhi, while Xiang, Gan, and especially Yue show only limited tonal alternations. Furthermore, sandhi processes take different forms in different dialect (sub)groups: tone deletion and tone spread, widely attested in Wu, are all but unknown among Mandarin and Min dialects.[2] It has been often noted that while northern Wu has a left-prominent prosodic structure, Min, Mandarin, and southern Wu exhibit a right prominence. This difference in rhythmic organization entails far-reaching consequences in tone sandhi behavior. As these and other generalizations hold across groups of dialects, it is often useful to identify the group membership of a particular dialect under discussion.

While Yuan (1960) still serves as a standard reference and most informative overview of Chinese dialectology, more recent surveys can be found in Egerod (1967), Norman (1988), and You (1992). Intensive research in the genetic classification and geographical distribution over the last two decades or so has culminated in *Language Atlas of China* (1987). The *Atlas* divides Chinese dialects into ten groups as shown in table 1.1. Mandarin, spoken by roughly 65 percent of the entire population of China, covers the largest area – basically the entire region north of the Yangzi river and the southwestern provinces (Yunnan, Guangxi, Guizhou, Sichuan). The compilers of the *Atlas* have separated the dialects spoken in Shanxi and adjacent regions of Hebei and Shaanxi from the surrounding Mandarin dialects, and put them under the Jin group. The remaining eight groups – sometimes collectively known as the "southern" dialects – are all concentrated in the southeastern corner. Aside from its main "homeland" located at the borderland where Jiangxi, Fujian, and Guangdong meet, pockets of Hakka speakers are found in Guangxi, western Guangdong, Taiwan, and

[2] That is, outside of the well-known so-called "neutral tone" phenomena.

Table 1.1. *Chinese dialects*

Group	Speakers (in millions)	Location (Provinces)	Representative Dialects
Mandarin	662.2	north of Yangzi rivers, and south-west provinces	Beijing, Tianjin, Ruicheng
Jin	45.7	Shanxi, north Shaanxi, west Hebei	Pingyao, Changzhi
Wu	69.8	south Jiangsu, Zhejiang, south-east Anhui	Shanghai, Suzhou, Danyang, Chongming, Zhenhai, Tangxi, Wenzhou, Wenling
Hui	3.1	south-east Anhui, west Zhejiang	Tunxi
Gan	31.3	Jiangxi, east Hunan	Nanchang
Xiang	30.9	Hunan	Changsha
Min	55.1	Fujian, Taiwan, east Guangdong, Hainan (south-east Asia)	Fuzhou, Xiamen, Chaozhou, Taiwanese, Wenchang
Yue	40.2	Guangdong, east Guangxi (south-east Asia, Americas)	Cantonese, Taishan
Pinghua	2.0	south Guangxi	Nanning
Hakka	35.0	south Jiangxi, west Fujian, east Guangdong, parts of Taiwan	Meixian, Changting, Pingdong

scattered over a large area of Sichuan. Even more far-flung is the Min (super)group. Specifically, varieties of southern Min are spoken not only on the mainland (Fujian and eastern Guangdong), but have spread over much of the islands of Taiwan and Hainan, and the Leizhou peninsula in southwestern Guangdong.

Citing *Renmin Ribao* (Overseas edition, March 11, 1989), R. Li (1989: 164) estimates overseas Chinese population at somewhere between 26.8

and 27.5 million, most of them living in southeast Asia (25 million) and the Americas (1.8 million), with the rest being scattered over Europe (380,000), Oceania (180,000) and Africa (80,000). Most of the overseas Chinese speak one or other Chinese dialect, in descending order of numerical strength: Yue (including Cantonese, 11 million), Southern Min (including Xiamen, Chaozhou, 8.6 million), Mandarin (3.5 million) and Hakka (0.75 million).

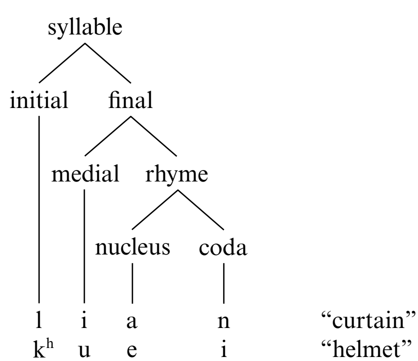
2 Historical background

For reasons that will become apparent (see section 5), an elementary knowledge of historical Chinese phonology is indispensable as a background to the understanding not only of tone sandhi, but of tonal systems in modern dialects. Not only do traditional tonal categories provide us with a handy common frame of reference that holds relatively constant across dialects, but in some cases tone sandhi rules are unstatable without reference to the same classical categories (see section 5.1).

2.1 Middle Chinese tonal categories

We start with the syllable. In most Chinese dialects, the maximal syllable consists of CGVX, that is, a consonantal onset, a prenuclear onglide, the nucleus, and a coda (which can be either an offglide, a nasal, or a voiceless stop). The hierarchical structure of the syllable is captured by the following diagram, labeled with the traditional terms commonly used in Chinese philological literature:

(1)



The "initial" is, of course, the onset. The "final" includes the medial (onglide) as well as the rhyme. The rhyme consists of a nucleus and a coda, which can be either an offglide or a consonantal ending.[3] The nucleus is the only obligatory syllabic element: thus bare vowels [i] "to heal," [u] "house," [ü] "rain," [a] (a prefix), [ə] "hungry" etc. all constitute legitimate syllables in Standard Mandarin. There remains some ambiguity regarding the status of the medial: whether the prenuclear glide belongs with the onset, or forms part of the final. For the on-going debate regarding subsyllabic constituency, I refer the reader to Lin (1989), Duanmu (1990a), and Bao (1990b, 1996a).

For our purposes, it suffices to note that syllables fall into two classes: (i) "**checked**" **syllables**, namely syllables ending in an occlusive coda (-p,t,k, often reduced to a glottal stop -q); (ii) "**smooth**" or "**slack**" **syllables**, namely either an open syllable CV (possibly with an offglide), or a syllable closed by a nasal stop. This dichotomy, whose tonological significance will become transparent immediately below, is quite robust not only in Chinese, but across other tone languages of southeast Asia, notably Kam-Tai and Miao-Yao, where the two syllable types are known by more colorful and expressive terms such as ***staccato*** vs. ***legato***, or ***dead*** vs. ***live*** syllables (cf. Gandour 1974, M. Hashimoto 1984, Thongkum 1987, Thurgood 1992). For short, I will sometimes use CVq and CVN to symbolize these two types of syllables.

Four tonal categories, referred to by their traditional nomenclature as *ping*, *shang*, *qu*, and *ru*, have been firmly established since Middle Chinese (hereafter MC; approximately from AD 200 to 900), as reflected in the pronouncing dictionary *Qieyun* (AD 601), a landmark in the history of Chinese phonology.

(2) Middle Chinese tone categories

	traditional name	gloss
I	*ping*	"level"
II	*shang*	"rising"
III	*qu*	"departing"
IV	*ru*	"entering"

Tone IV occurs exclusively with checked syllables, while tones I, II, and III are associated with smooth syllables. This cross-classification of tones

[3] Sometimes both. Thus, the northern Min dialect of Fuzhou has in its syllable inventory words like [souŋ] "sour," [keiq] "orange," etc. (-q = glottal stop).

and syllable types is motivated by the observation that checked syllables tend to have an impoverished tonal inventory, and exhibit markedly different sandhi behavior compared to smooth syllables, as will become amply evident in the ensuing chapters. For this reason, Chinese linguists often talk about "**smooth tones**" (*shu sheng* = tone I, II, III) and "**checked tones**" (*ru sheng* = tone IV). Throughout this book, I will suffix the symbol -q to the tone letters representing a "checked tone," while leaving the "smooth tones" unmarked. Thus, 55q, 13q etc. stand for a high level and low rising tone linked to a checked syllable CVq.

Based chiefly on the Japanese monk Annen's description in *Xi-tan zang* (or *Shittan zô*; written in AD 880), Mei (1970:109–110) reconstructs the following tonal values for MC (around 8th century):

(3) Middle Chinese tone values

	categories	reconstructed phonetic values
I	Level (*ping*)	long, level, and low (with two allotones)
II	Rising (*shang*)	short, level, and high
III	Departing (*qu*)	longish, probably high and rising
IV	Entering (*ru*)	short, with uncertain pitch and contour

Some of the descriptive terms for pitch height and contour are taken directly from Annen, who characterized tone I and II as "*zhi di*" (straight and low) and "*zhi ang*" (straight and high).[4] The hypothetical durational distinction is based primarily on the ancient buddhist practice of using tone II and III syllables to transcribe Sanskrit short and long vowels, respectively. As noted above, entering tone syllables end in an oral stop -p,t,k. Not surprisingly, Annen describes the entering tone as "*jing zhi*" (abruptly stop), a "checked" quality that is still readily observable in those modern dialects that have preserved the old p,t,k codas (often reduced to a glottal stop, hereafter symbolized as -q). The reconstruction of tone III is somewhat more speculative: it is inferred from the fact that tone II syllables with a voiced obstruent onset had merged with tone III.[5] Since merger presupposes a certain phonetic affinity, and since tone II is

4 Pulleyblank (1978:178) interprets *zhi ang* as "straight rising" instead, citing as evidence another contemporaneous document *Yuanhe Yunpu* (806–827), in which tone II is described as "*li er ju*" (stern and rising), where *ju* (lit. to lift up) clearly denotes a rising pitch movement.

5 This is clearly indicated in Annen's statement that tone II only occurs with *qing* "light" syllables, i.e. syllables with voiceless (and sonorant) initials.

known to be high, it stands to reason to assume that tone III also had a high pitch at the time the merger took place.[6]

Hirayama (1974, 1975) and Ting (1984) have made attempts at reconstructing the tonal values of Proto-Min and Proto-Wu, respectively.

Our thumbnail sketch of ancient Chinese tonology would not be complete without an aside on the hypothesis first put forward by Haudricourt (1954a, b, 1961), now generally referred to by the broader term "tonogenesis."[7] Haudricourt advanced the theory that the archaic Chinese tonal system arose through the loss of certain final consonants, in an evolution that parallels Vietnamese. Specifically, Haudricourt maintains that Archaic Chinese tone II, III and IV originated from CVq, CVs and CVk, respectively (-q represents a glottal stop, -s is a sibilant, and -k stands for any of the full oral stops -p,t,k). Crucially, Pulleyblank (1978) extends this hypothesis down to the more recent historical period of Middle Chinese. Specifically, he claims that the so-called "tones" actually corresponded to different syllable types prevailing in Late Middle Chinese (8th century), which still retained the old consonantal desinences. The Haudricourt–Pulleyblank hypothesis has found both supporters (Mei 1970, Sagart 1986) and skeptics (Ting 1981, 1996, Ballard 1985, 1988).

2.2 Tone split

The four Middle Chinese tones have undergone various splits and mergers. Tone split is sensitive to various phonological conditions, most notably the voicing contrast in the syllable onset,[8] as illustrated by the northern Wu dialect of **Songjiang** which, like all other Wu dialects, still maintains the voiced/voiceless contrast in the onset. Each of the MC tonal categories is split neatly into a high and a low register – known in traditional terminology as *yin* and *yang* – yielding a perfectly symmetrical eight-tone system (data from *Jiangsusheng he Shanghaishi Fangyan Gaikuang* 1960). In each

[6] For comparison, here is Ting's (1996:152) reconstruction of the MC tonal values, based on *Xi-tan zang* and other evidence (including comparative):

I	Level (ping):	level, probably low
II	Rising (shang):	high-rising
III	Departing (qu):	falling, probably mid-falling
IV	Entering (ru):	abrupt and short

Ting rejects any length contrast among tones I, II, and III.

[7] Coined by Matisoff (1970, 1973); see Hombert (1975, 1978), Hombert, Ewan, and Ohala (1979), Mazaudon (1977), and references cited therein. Utsat, a Chamic Austronesian language spoken on the Hainan island, instantiates a particularly transparent case of transition from an atonal to a fullblown tonal system. See Thurgood (1992) for details.

[8] Yip argues at length that what is crucial in conditioning tone split is not [voicing] but [murmur]. For details, I refer the reader to Yip (1980:211–242, and 1993b:249–254).

case, the **yang** register with a voiced onset has a lower pitch value than the corresponding **yin** register. Songjiang is fairly typical of Wu dialects.

(4) Songjiang register split

register		tone			
		I	II	III	IV
a.	high (*yin*)	53	44	35	5q
b.	low (*yang*)	31	22	13	3q

5q, 3q indicate checked tones

The pitch values are indicated by the familiar tone digits first introduced by Y-R. Chao (1930). The tonal space is idealized as a five-point vertical scale, where 5 and 1 represent the highest and the lowest pitch respectively.[9] Thus 53 and 13 stand for a high-falling and a low-rising tone respectively. Examples are given below:

(5)	I	a.	*ti*	53	“low”
		b.	*di*	31	“lift’
	II	a.	*ti*	44	“bottom”
		b.	*di*	22	“younger brother”
	III	a.	*ti*	35	“emperor”
		b.	*di*	13	“field”
	IV	a.	*paq*	5q	“hundred”
		b.	*baq*	3q	“white”

Needless to say, the voice-sensitive split into two registers is not always as neat or symmetrical. Take **Beijing Mandarin**. The correspondence between the MC tonal categories and their modern phonetic values is summarized in the following table (based on Chen 1976:152).

(6) Beijing Mandarin

MC onset	MC tones			
	I	II	III	IV
voiceless	55	213	51	55, 35, 213, 51
sonorant	35			51
voiced obstruent		51		35

The leftmost column indicates the three types of Middle Chinese onsets: voiced and voiceless obstruents, and sonorants (including liquids, nasals,

9 Exactly the opposite of the convention that prevails in African and Amerindian tonological literature.

and ∅-initials). Notice that tone I splits along the familiar [±voiced] division. Tone III remains a single cohesive category. Tone II also bifurcates along the voicing line, but in this case sonorants side with the voiceless rather than the voiced obstruents. Furthermore, the voiced obstruent onset syllables that split off from tone II have merged with tone III syllables. Finally, tone IV words (originally associated with checked syllables) are redistributed among the other tonal categories, conditioned by the three-way contrast between voiceless, sonorant, and voiced obstruent initials.[10] Note that two sweeping historical changes have occurred in Beijing Mandarin: all voiced obstruents have become voiceless, and all checked syllables (CVq) have lost their stop endings entirely. This means that both voicing and smooth vs. checked syllable contrasts are now recoded in purely tonal terms. In short, the evolution from MC to the tonal system of Beijing as we know it today entails the following historical processes:

(7) a. Register split of tone I
  b. Tone IIb merges with tone III
  c. Redistribution of tone IV among other tonal categories
  d. Devoicing and, in some cases, aspiration of voiced obstruents
  e. Loss of obstruent codas

Needless to say, (a, b, c) must precede (d) since the former are voice-sensitive, a distinction that is neutralized by the latter. Furthermore, (c) must pre-date (e) since the defining characteristic of tone IV is CVq, with an oral stop coda (symbolized by -q), which has dropped out via (e). Some examples follow:

(8)	MC	Standard Mandarin		
I	*tang*	*tang*	55	"ought to"
	lang	*lang*	35	"wolf"
	dang	*t'ang*	35	"sugar"
II	*tang*	*tang*	214	"party"
	lang	*lang*	214	"bright"
	dang	*tang*	51	"to swing, sway"
III	*tang*	*tang*	51	"to pawn"
	lang	*lang*	51	"wave"
	dang	*tang*	51	"to procrastinate"
IV	*t'ak*	*t'uo*	55	"to entrust"
	lak	*luo*	51	"to fall"
	dak	*tuo*	35	"to stroll, pace"

[10] Tone IV words with a voiceless initial are scattered unpredictably among all four tonal categories in modern Beijing.

Linguists have long noted the pitch-depressing effect of voiced obstruents, and sought to explain the cross-linguistic patterns in physiological terms.[11] For a general discussion and critical review see Hombert (1978), Hombert, Ohala, and Ewan (1979), and references cited there. For our purposes, it should be noted that subsequent historical changes – in particular, devoicing – may intersect and obscure the phonetically motivated partition of tonal categories into a high and a low register. It is not uncommon for the *yang* or b-register to show a high tone in a modern dialect instead of the expected low register, in a process sometimes referred to as "register flip-flop." A. Hashimoto (1986) sampled 997 dialects, and found 340 cases of register reversal. For this reason, we will often simply refer to register a (*yin*) and b (*yang*), to dissociate the relatively constant tonal categories from their often unpredictable phonetic values.

Somewhat less well known, but nonetheless quite common among Sino-Tibetan languages, is tone split along the dividing line between plain and aspirated onsets. Ye (1983) reports a three-way split of MC tones resulting in a perfectly symmetrical twelve-tone pattern in the Songling variety of **Wujiang**, also a northern Wu dialect.

(9) Wujiang three-way tone split

		I	II	III	IV
voiceless	plain	55	51	412	5q
	aspirated	33	42	312	3q
voiced[12]		13	31	212	2q

Since Wujiang has retained voicing and aspiration, the multiple splits merely produce allotonic variations rather than giving rise to new tonal categories. Examples illustrating the allotonic distribution within the four MC tonal categories follow:

(10)	I	55	*tiɪ*	"fall, topple"
		33	*t'iɪ*	"day, sky"
		13	*diəu*	"head"
	II	51	*tø*	"short"
		42	*t'i*	"body"
		31	*dᴇ*	"light, insipid"

[11] Related is the blocking effect of voiced consonants on H-spread. Conversely, voiceless consonants tend to block L-spread (see Hyman and Schuh 1974).

[12] Including sonorants and ∅ or vocalic onset.

III	412	*tE*	"toward"
	312	*t'E*	"to withdraw"
	212	*dəu*	"big"
IV	5q	*tiəq*	"drop"
	3q	*t'iəq*	"iron"
	2q	*doq*	"to read"

Aspiration-triggered tone split has been widely attested both within the Sinitic family and beyond (cf. Ho 1989). Shi (1994) cites no less than 111 languages within the Sino-Tibetan phylum that instantiate such a tonal development. In addition to 23 and 22 instances in the Gan and Wu groups of Chinese, there are 66 other cases: 44 in Dong (belonging to the Kam-Tai branch), and 22 in the various Miao-Yao languages, both considered by some to be part of the Sino-Tibetan family. A plausible phonetic explanation is offered in Hombert et al. (1979) and Shi (1994).

Tone splits along other phonological parameters have been reported in the literature: breathy voice, prenasalization, fortis vs. lenis consonants, vowel height, length and tensity, etc. (Hombert 1978, Endô 1994). Surprisingly, despite the well-known intrinsic pitch variations associated with vowel height,[13] tone split along the high/low vowel distinction is so rare that Hombert et al. (1979:52) state flatly: "It would seem that the interaction between tones and vowel height works in only one direction: tone can affect vowel height, but not vice-versa."[14]

2.3 Tone mergers

MC tones have undergone two sweeping mergers. First, tone IIb has splintered off from IIa, and falls together with tone III across all dialect groups, suggesting an early onset of this sound change. The exact membership of IIb is defined somewhat differently from dialect to dialect, owing to the "amphibious" nature of the sonorants. Specifically, the sonorants behave sometimes as voiced obstruents, sometimes as voiceless (more

[13] Attested in tonal (Itsekiri, Yoruba, Ewe) as well as non-tonal languages (English, Danish, German, Japanese, Korean, French, Serbo-Croatian, Hungarian); see Beckman (1986:129) for sources and references. Similar data on Chinese can be found in Wu and Cao (1979).

[14] Fuzhou instantiates a classic case of tone-on-vowel influence: a low-to-high tone change in sandhi contexts induces a concomitant vowel change:

ei	→	i
ou		u
øy		y
ɔi		øy
ai		ei
au		ou

precisely, as unmarked for voicing). This somewhat complicated picture is summed up in the following diagram, where p, m, and b stand for the three classes of initial consonants: voiceless obstruents, sonorants, and voiced obstruents:[15]

(11)

p	m	b	
IIa	IIb	IIb	Suzhou, Changsha, Guangzhou, Fuzhou
IIa	IIa	IIb	Nanchang, Shuangfeng, Xiamen, all northern dialects

However defined, tone IIb has merged with tone III in most dialects – with the notable exception of Wenzhou (Wu), Guangzhou (Yue) and Chaozhou (Min), which have kept IIb apart as a distinct tonal category.

The second major merger stemmed from the weakening and loss of MC -p,t,k endings, marked by the entering tone IV. We can distinguish four stages of this development that must have taken centuries to run its course.

a. The original state of affairs is still visible in many Yue and southern Min dialects, which have preserved a full-fledged series of obstruent codas.
b. The occlusive endings -p,t,k have weakened into an undifferentiated glottal stop -q in most Wu dialects.
c. In yet others, this -q has dropped out altogether. However, the original CVq syllables have stuck together as one cohesive tonal category. This is the case with Changsha (of the Xiang group), which no longer has CVq syllables, but has maintained a separate class identifiable by means of a distinctive tone contour [24].
d. Finally, the stop endings have disappeared without a trace, segmentally or otherwise.

For most purposes, the only relevant distinction is between checked and smooth tones, regardless of the degree to which the original occlusive endings have been preserved or reduced. For this reason, in the rest of this book I will distinguish only between T and Tq (where T stands for

[15] This table is derived from the computerized dialectological corpus DOC [Dictionary on Computer], based on the first edition of *Hanyu Fangyin Zihui* (Beijing University, 1962), which did not include Yangjiang and Jian'ou. For a recent description of DOC, see Cheng (1994a).

For a plausible historical account of the "amphibious" behavior of sonorants with respect to tone splits and mergers, see Yip (1980:240f.).

any tone). Thus 5q means that the high level tone corresponds to an MC tone IV, originally associated with the CVq syllable type.

What is striking about the evolution of MC tones is the stability of categorical membership in the face of phonetic diversity. In other words, tone A in dialect X corresponds with remarkable regularity with tone B in dialect Y, regardless of the phonetic shapes of tones A and B. Take the modern reflexes of tone Ib. It has seven distinct phonetic shapes {35 : 42 : 24 : 213 : 31 : 55 : 34} in the seven Mandarin dialects represented in *Hanyu Fangyin Zihui* (second edition, Beijing 1989). The inter-dialectal correspondence is as categorically systematic as it is phonetically heterogeneous. How a single subcategory tone Ib evolved into such a wildly disparate set of modern reflexes in closely related languages is still poorly understood. For an overview of the complex diachronic developments of MC tones across the full spectrum of Chinese dialects, see the large scale study conducted by Chang (1975), based on all the sources available at the time.

3 Tone patterns in present day dialects

The historical processes of splits and mergers have given rise to a wide variety of tonal systems in the modern dialects of Chinese, ranging from three to ten tonal categories, according to Cheng (1973b:96). Synchronic tone patterns are traditionally described by specifying two parameters, namely pitch *height* (high, low, mid, etc.) and tone *shape* (even, rising, falling, falling-rising, or rising-falling).[16] Take the **Songjiang** case alluded to in section 2.2 for illustration (data from *Jiangsusheng he Shanghaishi Fangyan Gaikuang* [Nanjing, 1960], p. 11).

(12) Songjiang tone system:

σ type	onset	even	rising	falling
CVN	voiceless	44	35	53
	voiced	22	13	31
CVq	voiceless	5q		
	voiced	3q		

"Voiced" includes both voiced obstruents and sonorants

16 Falling-rising and rising-falling are also known as "concave" and "convex" tones respectively after Wang (1967).

It has eight phonetic tone shapes. However, since Songjinang, like most other Wu dialects, has preserved the voiced/voiceless onset as well as the CVq/CVN contrast, the eight tones can be reduced to three contrastive categories, differentiated only in terms of tone shape (level, rising, falling), with predictable pitch height. Thus, the level tone has four allotones {44, 22, 5q, 3q} in phonetically definable complementary distribution. However, the prevailing practice is to treat the checked tones as if they constituted distinctive categories apart from their smooth counterparts, not without some justification, in view of the peculiar sandhi behavior and distributional restrictions of checked tones. Note in passing that Songjiang is also typical in restricting contour tones {35, 13, 53, 31} to smooth (CVN) syllables.

By a similar process we can reduce the 12 tones of Wujiang (see section 2.3) neatly into a three-tone system, by exploiting the redundancies implicit in the two-way contrast between CVN and CVq on the one hand, and the three-way opposition between voiced, plain voiceless, and voiceless aspirated onsets, on the other.

However, the picture is somewhat less symmetrical and clear-cut in some cases when we look more closely. **Shanghai**, a related but far better known Wu dialect, serves to illustrate the point. This dialect has five citation tones {53, 34, 23, 5q, 12q}, corresponding to Middle Chinese categories Ia, IIIa, b and IVa, b. Their distributions among syllables with the three classes of initial consonants are summarized in the following table, where "+" and "–" indicate the cooccurrence or its absence, of the relevant classes of tones and initials.

(13) Shanghai tone system

onsets	Ia	IIIa	IIIb	IVa	IVb
	53	34	23	5q	12q
voiceless[17]	+	+	–	+	–
sonorant	+	(+)	+	(+)	+
voiced obstruent	–	–	+	–	+

There is no question that the two high register tones [53] (MC tone Ia) and [34] (tone IIIa) stand in contrast, as exemplified by [tɔ 53] "knife" vs. [tɔ 34] "island" and countless other minimal pairs. On the other

[17] Including ∅-initial.

hand, it is possible to collapse the checked tones [12q] and [5q] with [23] and [34] respectively.[18] The question concerns whether [34] contrasts with [23] and [5q] with [12q]. A check of the syllabary in Xu et al. (1988:16–23, 78–100) yields an ambiguous answer: the high register tones [34, 5q] in principle occur only with voiceless initials, while the low register tones [23, 12q] only go with voiced onsets.[19] However, there are sporadic exceptions forming a few minimal doublets and triplets as exemplified below:

(14)	a.	*nu*	34	IIIa	"diligent"
			23	IIIb	"sweet rice"
	b.	*wAq*	5q	IVa	"to dig"
			12q	IVb	"slippery"
	c.	*mE*	53	Ia	"every"
			34	IIIa	"beautiful"
			23	IIIb	"slow"

The parenthesized (+) in (13) (shaded cells) is intended to signal the sporadic and exceptional nature of the cooccurrence of IIIa and IVa with voiced initials (sonorants). We may discount a handful of minimal pairs like those cited above, or rephonemicize the minimal pairs like "diligent" and "sweet rice" as /*nu* LH/ and /*nhu* LH/ respectively, where h stands for murmur or breathy voice. Likewise, "to dig" and "slippery" are given the lexical representations /*waq* LH/ and /*whaq* LH/. The phonetic values of LH are then determined by the segmental composition of the tone-bearing syllables. This is the practice adopted, for instance, in Sherard (1972), and followed by Yip (1980, 1993b), Duanmu (1990a), and Jin (1995). Either way, we could in principle compress Shanghai into a two-tone system, where the rising tone has four allotones, predictable (by and large) on the basis of voicing and syllable type:

[18] In principle, [5q] could go with the high falling [53] as well. However, most analysts group [5q] with the high rising [34] instead because they display a similar behavior in sandhi contexts.

[19] Which is actually a convenient oversimplification. It has been well known since Chao (1928, 1935) that voiced stops are voiced only in word-medial positions; in initial positions they are actually voiceless (by the standard test of voice onset time; cf. Lisker and Abramson 1967), contrasting with their voiceless counterparts only in terms of the closure time, longer for voiceless and shorter for voiced stops. For a critical survey of previous studies and original experimental data concerning voicing in Wu dialects, see Shen and Wang 1995.

(15)

		rising /LH/	falling /HL/
CVN	voiceless	34	53
	voiced	23	
CVq	voiceless	5q	
	voiced	12q	

Primary sources and secondary literature are not consistently explicit on the phonemic status of the various phonetic tones. Fortunately, it is not often the case that analysis crucially hinges on the phonemic status of the tones in question. In most cases, it suffices to take the citation tones (i.e. tones that appear with monosyllabic root morphemes pronounced in isolation) on their face value, and note their alternations in various sandhi contexts.

Cantonese (spoken in Hong Kong, Guangzhou, and by many overseas Chinese) has one of the more complex tonal systems, characteristic of the Yue group to which it belongs. The nine tones of Cantonese are customarily cross-classified as follows (based on Hashimoto 1972:92 and *Hanyu Fangyin Zihui*, 2nd ed. 1989):[20]

(16) Cantonese tone system

	level		rising	falling
	CVN	CVq		
high (*yin*)	33	5q; 3q	35	53 (~ 55)
low (*yang*)	22	2q	23	21

The high falling tone [53] has [55] as a free (stylistic) variant. Ancient voiced obstruents have undergone devoicing in Cantonese. As a consequence, *yin* and *yang* registers, originally derived from the voiced/voiceless dichotomy, now contrast exclusively in terms of pitch height. On the other hand, the stop endings (-p,t,k) are preserved intact. This means that [5q, 3q] and [2q] are in complementary distribution with [33] and [22]. Nevertheless, as noted before, the checked tones [5q, 3q, 2q] are treated as separate tonal categories in accordance with the descriptive conventions adopted by most dialectologists in China. Finally, the bifurcation of the *yin* register

[20] I adopt the tone values of *Hanyu Fangyin Zihui* in the interest of cross-dialect comparability.

checked tone into [5q] and [3q], attested in many other Yue dialects, is conditioned by the opposition between what A. Hashimoto (1972:176f.) refers to as lax (short) and tense (long) vowels, for instance [ɪ] vs. [ɛ:] and [a] vs. [a:].[21] The nine tones of Cantonese are illustrated by the following three sets of monosyllabic root morphemes:

(17) Cantonese

MC categories	tone values						
Ia	53 (~ 55)	si	"poetry"	fu	"husband"	wan	"warm"
Ib	21	si	"time"	fu	"to hold"	wan	"cloud"
IIa	35	si	"to send"	fu	"bitter"	wan	"to look for"
IIb	23	si	"market"	fu	"wife"	wan	"to allow"
IIIa	33	si	"to try"	fu	"rich"	wan	"to shut up"
IIIb	22	si	"affairs"	fu	"father"	wan	"to transport"
IVa-i	5q	sɪk	"to know"	fat	"sudden"	wat	"twisted"
IVa-ii	3q	sɛ:k	"lead"	fa:t	"law"	wa:t	"to dig"
IVb	2q	sɪk	"to eat"	fat	"to punish"	wat	"pit" (of fruit)

Bobai, also a Yue dialect (spoken in Guangxi) and one of the earliest tone languages reported in the classic work of Wang Li (1932), is often cited as the Chinese dialect endowed with the richest tonal repertoire, boasting ten tonal categories. Although Bobai has phonetically quite different tone shapes from Cantonese, in terms of tonal categories it differs from Cantonese only in that tone IVb is further split into two subcategories.[22]

Interpretation of phonetic data

From a broader perspective, Cheng (1973b) surveyed 737 contemporary Chinese dialects, and found 69 different phonetic shapes. The phonetic transcriptions call for judicious interpretation. With rare exceptions, phonetic transcriptions are based on aural judgment, and vary according to different practices and implicit assumptions on the part of the fieldworkers. Furthermore, the five-point pitch scale specifies a far greater number of tone shapes than one would ever need to describe any one language, as a consequence, forcing arbitrary choices upon the fieldworker in many cases. Take the tone shape [54] (attested in 57 dialects). One cannot tell a

[21] Hashimoto's (1972) inverted [a] : [ʌ:] correspond to Yuan's (1960) [a] : [a:]. For typographical simplicity, I adopt Yuan's transcription.
[22] The bipartition of tone IVb is also attested in Yangjiang, another Yue dialect.

priori whether it is basically a high level tone [55] with a slight declination effect, or a variant of [53], or for that matter, [454], and so forth. By the same token, if a dialect has only one rising tone, whether one transcribes it as [24], [34], [35] etc. depends as much on personal preferences and (implicit) theoretical assumptions as on the objective phonetic reality. For instance, if one partitions the pitch range into high (4–5) and low (1–2) registers – interpreting 3 as either high or low as the case may be – and if one disallows, for theoretical reasons, cross-register contour tones, one would reject [24] in favor of [34] or [35]. The analyst has to make judgment calls of this sort on a case-by-case basis.

One concrete example will suffice to illustrate the need for a judicious interpretation of the raw phonetic data. According to Norman (1973), **Jianyang** has the following citation tones:

(18) Jianyang citation tones

even	rising	falling
33	35	53
		43
		32
		31
		21

On the face of it, we have an oddly skewed system, with one even, one rising, but five falling tones. It not only displays a lopsided preponderance of contour tones, but is highly marked, if not impossible given a classificatory scheme like Yip's (1980, 1989) or Bao's (1990a), which allows only a maximum of four level tones, two rising and two falling tones. I agree with Yip (1980:206) who suggests that in all likelihood, three of the five falling tones, namely [43, 32, 21], ought to be considered basically level tones [44, 22, 11] with a slight downdrift or final fall in pitch, a predictable phonetic effect noted in the tonological and intonational literature.[23] On this view, the Jianyang system should be revised as follows:

[23] On the phenomena of "declination" and boundary low cf. Pierrehumbert (1980), Liberman and Pierrehumbert (1984), and, with particular reference to Chinese, Tseng (1981) and Shih (1988, 1991). Tseng (1981:143) notes, for instance, that the terminal fall is a special case of breath-group intonation.

(19) Jianyang tone system

even	rising	falling
44	35	53
33		
22		31
11		

In interpreting the raw phonetic data, we need to keep the overall sound pattern in mind and, where appropriate, follow the eminently sensible heuristic principles proposed by Maddieson (1978b:45f.) in discounting non-distinctive phonetic details. One common non-distinctive feature is the final lowering already noted above in relation to Jianyang. To cite another example, one may judiciously disregard certain finer details resulting from the undershooting of phonological targets. Take the two pitch curves in **Standard Thai** which we annotate as [214] and [451] in Chao's tone digits. Maddieson interprets them as basically rising [14] and falling [51] tones in Standard Thai, with two and not three tonal targets each; in his view, the initial portions of the pitch curves transcribed as [2] and [4] represent nothing more than imprecise approximations toward the intended targets [1] and [5].

4 Tones in context

Rich and highly developed as tonal systems have become in Chinese, they are surpassed in many instances by even more complex and intricate sandhi processes, which often drastically alter the phonetic shape of adjacent tones, when they come into contact with each other in connected speech. This tonal alternation in connected speech is what has been referred to as **tone sandhi**, and constitutes the subject matter of our inquiry.

Before we delve into the specifics of tone sandhi, it is useful to place tone sandhi in the context of the various types of tonal modifications. Tone shapes may be subject to the influence not only of the neighboring tones, but also of the overall intonation. Furthermore, morphotonemic alternations may be triggered not only by strictly phonological environments, but also by morphological contexts.