CAMBRIDGE STUDIES
IN LINGUISTICS 92

Tone Sandhi

Patterns Across Chinese Dialects

MATTHEW Y. CHEN

CAMBRIDGE

more information - www.cambridge.org/0521652723

This page intentionally left blank

Matthew Chen's landmark study offers the most comprehensive analysis to date of the rich and complex patterns of tone used in Chinese languages. Chinese has a wide repertoire of tones which undergo often surprising changes when they are connected in speech flow. The term tone sandhi refers to this tonal alternation. Chen examines tone sandhi phenomena in detail across a variety of Chinese dialects. He explores a range of important theoretical issues such as the nature of tonal representation, the relation of tone to accent, the prosodic domain of sandhi rules, and the interface between syntax and phonology. His book is the culmination of a ten-year research project and offers a wealth of empirical data not previously accessible to linguists. Extensive references and a bibliography on tone sandhi complete this invaluable resource which will be welcomed as a standard reference on Chinese tone.

Having taught for many years at the University of California, San Diego, MATTHEW CHEN is Professor and chair of Linguistics and Dean of the Humanities and Social Sciences faculty at the City University of Hong Kong. He has written numerous articles on linguistics and is Associate Editor of the *Journal of Chinese Linguistics*, University of California, Berkeley.

In this series

52 MICHAEL S. ROCHEMONT and PETER W. CULICOVER: *English focus constructions and the theory of grammar*
53 PHILIP CARR: *Linguistic realities: an autonomist metatheory for the generative enterprise*
54 EVE SWEETSER: *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure*
55 REGINA BLASS: *Relevance relations in discourse: a study with special reference to Sissala*
56 ANDREW CHESTERMAN: *On definiteness: a study with special reference to English and Finnish*
57 ALLESSANDRA GIORGI and GUISEPPE LONGOBARDI: *The syntax of Noun Phrases: configuration, parameters and empty categories*
58 MONIK CHARETTE: *Conditions on phonological government*
59 M.H. KLAIMAN: *Grammatical voice*
60 SARAH M.B. FAGAN: *The syntax and semantics of middle constructions: a study with special reference to German*
61 ANJUM P. SALEEMI: *Universal grammar and language learnability*
62 STEPHEN R. ANDERSON: *A-Morphous morphology*
63 LESLEY STIRLING: *Switch reference and discourse representation*
64 HENK J. VERKUYL: *A theory of aspectuality: the interaction between temporal and atemporal structure*
65 EVE V. CLARK: *The lexicon in acquisition*
66 ANTHONY R. WARNER: *English auxiliaries: structure and history*
67 P.H. MATTHEWS: *Grammatical theory in the United States from Bloomfield to Chomsky*
68 LJILJANA PROGOVAC: *Negative and positive polarity: a binding approach*
69 R.M.W. DIXON: *Ergativity*
70 YAN HUANG: *The syntax and pragmatics of anaphora*
71 KNUD LAMBRECHT: *Information structure and sentence form: topic, focus, and the mental representations of discourse referents*
72 LUIGI BURZIO: *Principles of English stress*
73 JOHN A. HAWKINS: *A performance theory of order and constituency*
74 ALICE C. HARRIS and LYLE CAMPBELL: *Historical syntax in cross-linguistic perspective*
75 LILIANE HAEGEMAN: *The syntax of negation*
76 PAUL GORRELL: *Syntax and parsing*
77 GUGLIELMO CINQUE: *Italian syntax and universal grammar*
78 HENRY SMITH: *Restrictiveness in case theory*
79 D. ROBERT LADD: *Intonational phonology*
80 ANDREA MORO: *The raising of predicates: predicative noun phrases and the theory of clause structure*
81 ROGER LASS: *Historical linguistics and language change*
82 JOHN M. ANDERSON: *A notional theory of syntactic categories*
83 BERND HEINE: *Possession: cognitive sources, forces and grammaticalization*
84 NOMI ERTESCHIK-SHIR: *The dynamics of focus structure*
85 JOHN COLEMAN: *Phonological representations: their names, forms and powers*
86 CHRISTINA Y. BETHIN: *Slavic prosody: language change and phonological theory*
87 BARBARA DANCYGIER: *Conditionals and prediction: time, knowledge and causation in conditional constructions*
88 CLAIRE LEFEBVRE: *Creole genesis and the acquisition of grammar: the case of Haitian Creole*
89 HEINZ GIEGERICH: *Lexical strata in English: morphological causes, phonological effects*
90 A.M.S. MCMAHON: *Lexical phonology and the history of English*
91 KEREN RICE: *Morpheme order and semantic scope: word formation and the Athapaskan verb*
92 MATTHEW Y. CHEN: *Tone sandhi: patterns across Chinese dialects*

Supplementary volumes

LILIANE HAEGEMAN: *Theory and description in generative syntax: a case study in West Flemish*
A.E. BACKHOUSE: *The lexical field of taste: a semantic study of Japanese taste terms*
NIKOLAUS RITT: *Quantity adjustment: vowel lengthening and shortening in early Middle English*

Earlier issues not listed are also available

CAMBRIDGE STUDIES IN LINGUISTICS

Tone Sandhi

To PL, sine qua non

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
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published in printed format 2000

ISBN 0-511-03307-9 eBook (Adobe Reader)
ISBN 0-521-65272-3 hardback

Contents

Preface

Although applied to the description of languages as diverse as Ewe of Ghana (Clements 1978), Mixtec of Mexico (Hunter and Pike 1969), and Kairi of Papua New Guinea (Newman-Petterson 1990), the term **tone sandhi** refers, first and foremost, to the kind of tonal alternations one typically finds in languages spoken in China and surrounding areas of the southeast Asian mainland. These languages share certain areal characteristics, notably highly developed tonal repertoires matched, in many cases, by even more complex tonal alternations in connected speech. The Sanskrit root of *sandhi* means junction, connection, combination, or liaison (cf. Allen 1962, Andersen 1986). Tone sandhi, therefore, *sensu stricto* describes phonetically conditioned morphotonemic alternations at the juncture of words or morphemes. Over time, however, tone sandhi has been extended to cover a number of related phenomena, including allotonic variations, intonational effects, and morphologically or syntactically conditioned tone changes. Such a broad and loose usage of the term has one advantage over any strict definition, as Andersen (1986:2) argues in a different context, in that it gives us a convenient label without prejudging the issues involved.

Even though the tonal/registral systems and their historical developments have been investigated extensively in such language groups as Tibeto-Burman, Miao-Yao, Kam-Tai, and Mon-Khmer, tone sandhi has not received as much attention[1] – with the notable exception of the Sinitic group. As a consequence, our study of tone sandhi has a distinctly Chinese focus and flavor, although I will not shy away from drawing on other languages for facts and inspiration.

[1] For instance the nearly 900-page report on the Kam-Tai languages spoken in China (J. Wang et al. 1984) devotes exactly two short paragraphs to tone sandhi in only one out of the eight languages surveyed, all of which have highly developed tone systems. It is not clear whether the Kam-Tai languages are tone-rich but sandhi-poor, or display sandhi phenomena that have gone underreported as in the early days of Chinese dialectology.

The investigation of tone sandhi has a long tradition in Chinese linguistics. The fourteenth-century pronouncing dictionary *Zhongyuan Yinyun* contains the earliest intimation of tone sandhi I am aware of. There the author, Zhou Deqing, noted that in versification "it would be best to avoid a pair of Rising tones or a pair of Departing tones." The Standard Mandarin tone sandhi rule as we know it today already found a clear formulation by the sixteenth-century Korean scholar Cui Shizhen, which I quote in part:

> If both syllables are in the Rising tone, then the circumstances make it difficult to retain the original tone. In this case, pronounce the first syllable like the voiced variety of the Level tone, and then the second syllable can retain its original tone when pronounced. (from Mei 1977:238f.)

A similar statement is found in the very first Western language grammar of Chinese written by Francisco de Varo, O.P. in 1682, whose statement is reproduced here:

> *mai*, en tercera tonada es comprar . . . mas juntandola con otra tercera, como *mai ping*, pronunciandolas juntamente el *mai* ba quasi a ser primera ["*mai*, in the third tone, means 'to buy' . . . but next to another third tone, as in *mai ping*, it nearly becomes the first tone when pronounced together"; tr. MYC] (Francisco de Varo, 1682, *Arte de la Lengua Mandarina*, p. 9)

In modern times, systematic descriptive work on Chinese dialects dates back to the 1920s. The early surveys tended to focus on the static tonal systems, and were sparse in information on the dynamic interaction of tones in context. Two early works by Chiu (1931) and Luo (1930) – both on the south Min dialect of Xiamen – are exceptional in their extensive treatment of sandhi phenomena and close attention to phonetic details. These are among the true pioneers of the study of tone sandhi. Since then important works, especially doctoral dissertations, have been devoted to this specialized topic. But it was the launching of the journal *Fangyan* in 1979 that marked the beginning of a dramatic explosion of empirical knowledge about the range and diversity of sandhi phenomena, some of which occur in obscure and hard-to-reach dialects. We get a glimpse of the overall picture from A. Hashimoto (1987 [1980]), Chen (1991a [= 1985]), and Ballard (1988). Since then our knowledge about the subject matter has broadened and deepened considerably, and the time is ripe for a new synthesis.

Two leitmotifs underscore the study of tone sandhi with particular reference to Chinese. The first concerns the internal structure of tone. Different hypotheses regarding tonal features and their geometrical arrangements make different predictions about the typology of possible tonal processes. Tone sandhi, therefore, serves as an effective diagnostic probe into the anatomy of the complex entity we call tone. The second recurrent theme of tone sandhi studies concerns the scope or domain of sandhi rules. Unlike most segmental phenomena, which tend to be localized,[2] tonal processes are notorious long-distance runners, sometimes spanning entire phrases and sentences. The precise definition of tone sandhi domains, therefore, raises intriguing questions about the interface between phonology and grammatical structure. There is a third, far less well-developed issue, namely the interplay among sandhi processes. Given a tonal string A–B–C, the sandhi form of the whole is typically the composite result of the elementary processes operating on the substrings A–B and B–C. How exactly the elementary processes interact to produce the ultimate sandhi output is a topic that has not been heretofore explored in depth.

Thematically, this book is organized as follows: after the introductory and stage-setting chapter, the book is divided into two parts. The first part, comprising chapters 2 to 6, deals with various types of tone sandhi phenomena. Chapter 2 focuses in particular on tonal geometry and the typology of sandhi rules. Chapters 3 and 4 are devoted to issues regarding how a sandhi process is implemented (e.g. directional iteration) and how one process may interact with another. These issues are of particular significance from the perspectives of Optimality Theory (Prince and Smolensky 1993, inter alia). Chapters 5 and 6 together constitute an in-depth investigation of one particular dialect, New Chongming. One of the startling discoveries is that this northern Wu dialect is well on its evolutionary path toward a classic accentual system.

The question of sandhi domains is taken up in part two, consisting of chapters 7 to 11. The scope within which tone sandhi rules operate ranges from a sublexical stress-foot to a phonological word, phonological phrase, and intonational phrase. One surprising finding is that the "Minimal Rhythmic Unit," which circumscribes the scope of tone sandhi in Beijing Mandarin (chapter 9), is not commensurate with any of the categories we know of conventional prosodic hierarchy. This and other findings are

[2] Vowel harmony and nasalization being two well-known exceptions.

summarized in the Concluding Remarks, followed by References and a Subject Index.

I take for granted not only generative phonology with its subtheories including autosegmental phonology, feature geometry, metrical phonology, and prosodic phonology, but also Optimality Theory, which by the time this book sees the light of day, will no doubt have become every practicing linguist's stock in trade. Where alternative descriptions of the facts are equivalent, I generally couch my analysis in conventional, generative terms. I will not hesitate to exploit the insights and formalisms of Optimality Theory when it sheds light on the issue at hand.

Much of the empirical data underpinning the present study is in the public domain, published in journals or being circulated through informal channels (unpublished doctoral dissertations, manuscripts, etc.). However, given the nature of the theoretical issues under investigation, published sources are often silent on critical aspects of the problem; as a consequence we had to develop new *kinds* of data, involving a variety of larger and more complex constructions. This is true even of relatively well-known and well-documented dialects, such as Standard Mandarin or Xiamen. I use myself as the principal informant for Xiamen and Mandarin, consulting other native speakers where subtle judgments may differ. In data sampling, I have resisted the butterfly collector's temptation, and opt for in-depth analysis of a few dialects selected for certain structural properties they serve to illuminate. I have collected a fair amount of original data on the following dialects: New Chongming and Old Chongming (northern Wu), Wenzhou (southern Wu), Tianjin (northern Mandarin), and Pingyao (Jin).

Acknowledgments

From the earliest conception to the final stages of redrafting this book, I have benefited from the inspiration and insights of countless friends and colleagues. But most of all, I am indebted to my past and current students who have nurtured and shared with me an abiding interest in the wondrous ways tones behave in connected speech. Chilin Shih, Tony Hung, Yuchau Hsiao, Hongming Zhang, and Huichuan Hsu and Lily Chan have all completed their doctoral theses on tone sandhi. They have shaped my thinking as much as I have theirs.

My thanks go to the informants who patiently sat through interminable sessions. In particular, I wish to name two informants who are linguists

in their own right: Pan Wuyun and Zhang Huiying. Pan Wuyun not only provided the raw data, but also offered some of the insights on which my analysis of Wenzhou (chapter 11) is based. Xu Baohua of Fudan University was my most gracious host in the fall of 1986 and the summer of 1990 during which most of the fieldwork on Wu dialects was carried out. Several of his colleagues, especially Pan Wuyun, Qian Nairong, Zhang Hongming, and Chen Zhongmin helped with various aspects of data elicitation and transcription. With patience and remarkable thoroughness, Lily Chan and Karanda Tang assisted me with the myriad minutiae of the final draft.

While the bulk of the present draft was written between 1994 and 1996 (during which the author was granted two quarters of sabbatical-in-residence), the underlying research took much longer. I have presented some of the preliminary results at various conferences and gatherings, including: a lecture series delivered at Centre de Recherches Linguistiques sur l'Asie Orientale (Paris, June 1990), talks delivered at the LSA Summer Institute (University of California, Santa Cruz, July 1991), Berkeley Linguistics Society Meeting (February 1992), Workshop on the Psychobiological Basis of Language (Taipei, December 1993), the Tilburg conference on "Derivational Residue" (Tilburg, October 1995), and on several occasions at East-Asian Linguistics Workshop (University of California Irvine, October 1993, October 1995), North-American Conference on Chinese Linguistics (Cornell University, May 1991; University of Delaware, May 1993; University of Southern California, May 1994; University of Wisconsin, June 1995; University of Illinois, May 1996), International Conference on Chinese Linguistics (Singapore, July 1992; Paris, June 1993), International Symposium on Chinese Languages and Linguistics (Taipei, July 1991, July 1992), Pan-Asiatic Linguistics Symposium (Chulalongkorn University, January 1992; Mahidol University, January 1996), and various colloquia (Berkeley, November 1990; November 1991; University of California, Los Angeles, April 1991; October 1995). I thank the organizers and the audiences of these meetings for their comments and criticisms.

An National Science Foundation grant (BNS-8608374, 1986–90) and a fellowship awarded by the Wang Institute of Graduate Studies (Tyngsboro, Mass., 1986–87) gave the initial impetus to this decade-long project by providing the time and wherewithal to make several trips to China and to bring informants to this country for extensive fieldwork. Over the years, the University of California, San Diego, Academic Senate has made smaller grants to pay for research assistance and related costs. In

grateful acknowledgment of their generous support, this book comes as a long overdue final report.

I wish to thank Richard Attiyeh, Vice-Chancellor of Research and Dean of Graduate Studies, University of California, San Diego, Joseph C-Y. Chen, Director of Wei-Kung Institute, and Yuchau Hsiao of National Chengchi University for their generous publication subsidy made to Cambridge University Press.

Notational conventions

Symbols	Meaning
T	Tone, tone root
T′	A modified or derived tone.
o	Zero tone, or unspecified for tone.
33, 24, 51 . . .	Tone letters devised by Y-R. Chao (1930), where digits indicate the pitch value on a five-point scale, 5 being highest. Thus 33, 24, 51 represent, a mid level, high rising, and a falling tone respectively.
H, M, L	High, mid, low. H, M, L and tone letters are used interchangeably throughout this book. Where appropriate, a dot separates tone digits and H, M, L, for instance to distinguish HM.L (= HM + L) from H.ML (= H + ML), or 3.52 (= 3 + 52) from 35.2 (= 35 + 2).
T1, T2 . . .	Tonal categories: tone 1, tone 2, etc. in a given system.
CVN	"Smooth" or "legato" syllable, i.e. ending in a vowel, offglide or a nasal.
CVq	"Checked" or "staccato" syllable, i.e. ending in an obstruent coda -p,t,k or a glottal stop -q.
MC	Middle Chinese
I, II, III, IV	MC tonal categories: *ping, shang, qu, ru*; tone IV cooccurs only with checked or staccato syllables CVq.
Xa,b . . .	Middle Chinese tonal category X, register *a* (typically high) or *b* (low). Middle Chinese Ia, Ib, IIa, IIb . . . are sometimes referred to as T1, T2, T3, T4 . . . , with odd / even numbers indicating register a and b respectively.
Tq	Checked tone, i.e. a tone associated with a checked syllable (CVq); thus 5q and 23q stand for a high and a low rising tone linked to CVq. Smooth or legato tones, i.e. tones linked to CVN are unmarked.

Symbols	Meaning
Hr, Lr	High register, low register, equivalent to [+upper] and [–upper]
h, l	Terminal tone segments, equivalent to [+raised] and [–raised] Lower case tone letters [h, m, l] are also used to indicate default pitch values H, M and L.
E, R, F	Even (= level), rising, falling
σ, μ	Syllable, mora
ϕ, p-word	Foot, prosodic word
p-phrase	Phonological phrase
IP	Intonational phrase
(x .), (. x)	Left-, right-prominent metrical unit
$C_1 \gg C_2$	Constraint 1 ranks above constraint 2
X } Y	Candidate X is more "harmonic" than candidate Y
☞	Picks out the winning candidate

Phonetic transcription

Examples are usually cited in the form they are found in the sources. Occasionally, the same pronunciation is transcribed differently by different authors. For instance, the Shanghai pronunciation of the word *Australia* is given as *qo.ta.li.ya* (Duanmu 1993a) and *ɔ.tʌ.li.iʌ* (Xu et al. 1988). I have made no systematic effort to standardize the phonetic transcriptions or phonemicizations across the board. The inconsistencies from one source to another are seldom crucial for our purpose; where they are relevant, they will be duly noted. Here are some recurrent symbols used here and their interpretations:

q	glottal stop
h̲	voiced / murmured h
y, w	glides of i,u
ü	front-rounded u
C′	aspirated
ng	velar nasal
ɤ	mid, back, unrounded vowel
E, I	(in small cap), lax vowels

Where phonetic transcriptions are not provided in the sources, I use the standard Pinyin system, which is also the accepted convention for transcribing Standard (Beijing) Mandarin. Pinyin differs from the IPA system chiefly in the symbols used for consonants. Here are the Pinyin symbols and their phonetic interpretation:

		labial	dental	retroflex	palatal	velar
stops	plain voiceless	b	d			g
	voiceless aspirated	p	t			k
affricates	plain voiceless		z	zh	j	
	voiceless aspirated		c	ch	q	
fricatives		f	s	sh	x	h
sonorants		m	n, l	r		ng

Occasionally when clarity demands, I separate syllables as well as tones by a dot. Thus *tian.an.men* "The Gate of Heavenly Peace" is syllabified as indicated (rather than *tia.nan.men*). More pertinently, MH.L = MH + L, while M.HL = M + HL; correspondingly: 35.1 = 35 + 1, but 3.51 = 3 + 51.

In citing examples, I often adopt the following format:

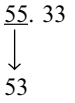
red	ten character	←	literal gloss
hong	[*shi zi*] "the Red Cross"	←	translation
(MH.LM)(LM)		←	input, base form
(MH. o) (LM)	Deletion		
(M. H) (LM)	Spread		
(M. H) (MLM)	M-Insertion	←	output, sandhi form

Here the semantic relation between "the Red Cross" and its component parts (lit. "the character for the word ten") is somewhat opaque.[3] I omit the gloss in those cases where the mapping between the literal (morpheme-by-morpheme) gloss and the English translation is transparent. The square brackets indicate morphosyntactic constituency, while the parentheses mark prosodic units, in this case the metrical feet. The sandhi rules that generate the intermediate or final outputs are named on the right by their labels given in the main text.

[3] The graphic representation of the root morpheme "ten" happens to be shaped like the cross: ✚.

Rule formalism

Occasionally I rotate the rewrite arrow clockwise by 90 degrees so that the target (which I underline for clarity) and the environment of a phonological rule would be right next to each other, rendering the rule more transparent. For instance:



is equivalent to

55 → 53 / ___ 33

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).

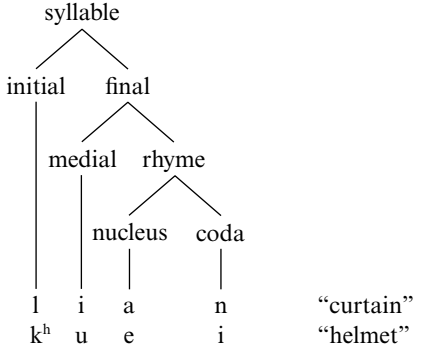
2 Historical background

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	*tʻiɪ*	"day, sky"
		13	*diəu*	"head"
	II	51	*tø*	"short"
		42	*tʻi*	"body"
		31	*dɛ*	"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	*tɛ*	"toward"
	312	*t'ɛ*	"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.

4.1 Tone sandhi

Beijing Mandarin exemplifies tone sandhi in its classical, canonical form. Recall that Beijing has a four-tone system illustrated by the oft-cited standard quadruplet:

(20)	T1	*ma*	[55]	"mother"
	T2	*ma*	[35]	"hemp"
	T3	*ma*	[214]	"horse"
	T4	*ma*	[51]	"to scold"

The rule in question turns a "dipping" i.e. falling-rising tone [214] into high rising [35] when followed by another [214]. Since [214] is traditionally labeled Tone 3, this processes is known as third tone sandhi, or **T3 Sandhi** for short. T3 Sandhi is illustrated by the following examples. (For clarity, a dot separates two tones represented by two to three tone digits.)

(21) a.	*xiao*		"small"
	214		
	xiao gou		"small dog, puppy"
	214. 214		base form
	35. 214		sandhi form
b.	*mai*		"buy"
	214		
	mai ma		"to buy a horse"
	214. 214		base form
	35. 214		sandhi form
c.	*mai*		"to bury"
	35		
	mai ma		"to bury a horse"
	35. 214		base form = sandhi form

Attention has been drawn to this phenomenon since the earliest descriptions of Chinese, presumably because this drastic change in tone shape is both perceptually salient and functionally significant – for instance, as a consequence of this systematic morphotonemic alternation, the expression "to buy a horse" becomes indistinguishable from "to bury a horse" (cf. Wang-Li 1967). But T3 Sandhi is but one of the many ways the idealized tone shape may be modified under the influence of neighboring tones and the overarching intonation pattern. Based largely on aural discrimination, Y-R. Chao (1968:26–29) mentions three others. First, [214] is obligatorily reduced to a simple low fall [21] everywhere except in utterance-final positions, where it may, but does not necessarily, show up

in its full form, with an upglide at the end. This means that, in conjunction with T3 Sandhi, tone 3 has three allotones, distributed as follows:

(22) Allotones of T3

		allotones		
		214	21	35
Sandhi contexts	before [214]	–	–	+
	utterance-final	+	+	–
	elsewhere	–	+	–

Here plus "+" and minus "–" indicate the occurrence or systematic absence of an allotone in a given context. Note that both the full [214] and the reduced [21] forms of T3 may occur in the utterance-final position; hence the notation "21(4)" in the examples that follow. Take the three occurrences of *xiao* "small," highlighted in bold face below.

(23) a. *dan* ***xiao*** "coward" (lit. gall + small)
214. 214 base form
35. 21(4) sandhi form

b. ***xiao*** *gou* "small dog, puppy"
214. 214 base form
35. 21(4) sandhi form

c. ***xiao*** *mao* "small cat, kitten"
214. 55 base form
21. 55 sandhi form

In utterance-final position, it shows up with the underlying [214], optionally reduced to [21] (case a); before another [214] it obligatorily turns into [35] (case b); elsewhere it may assume only the truncated form [21], also referred to as "half-third tone" (case c). Correspondingly, the rule responsible for this alternation is known as the **Half T3 Sandhi**.

Second, a rising tone [35] becomes a high level [55] when preceded by [55, 35] and followed by any tone. We shall have more to say about this rule (see chapter 7, section 2.3), generally referred to as **T2 Sandhi**, since [35] is the second tone in Standard Mandarin. A small handful of examples will suffice for our immediate purpose.

(24) a. *tian wen tai* "observatory" (lit. astronomy + platform)
55. 35. 35 base form
55. 55. 35 sandhi form

b.	*ren min bi*	"renminbi" (Chinese currency)
	35. 35. 51	base form
	35. 55. 51	sandhi form

There is a third rule, whereby neutral-toned syllables assume a pitch level that is determined by the preceding full tone. Specifically: neutral tone is predictably high [4] after [214], mid [3] after [35] and [55], and low [1] after [51]. This neutral tone rule – or **NT Sandhi** – is illustrated by the allotones of the inherently toneless enclitic *-de*, which functions, among other things, as the nominalizer:

(25)	a.	*zi de*	"the purple one"
		214. 4	
	b.	*hong de*	"the red one"
		35. 3	
	c.	*xin de*	"the new one"
		55. 3	
	d.	*da de*	"the big one"
		51. 1	

Finally, apart from the above phonologically determined sandhi processes, certain sandhi alternations are morphologically conditioned or lexically marked. One such alternation is referred to in Chang (1992:166f.) as the *Yi-bu-qi-ba* rule, because it applies specifically to a closed list of the four lexical items *yi* "one," *bu* "not," *qi* "seven" and *ba* "eight." This rule obligatorily changes the tone of these syllables into [35] when followed by another falling [53] tone syllable. As a consequence, we have a minimal pair such as this:[24]

(26)	a.	*bu dui*	"not correct"
		53.53	base tone
		35.53	*Yi-bu-qi-ba* rule
		cf. *bu hao* 53.214	"not good"
	b.	*bu dui*	"troops"
		53.53	no change

Bu "not" has the lexical tone /53/ as demonstrated by *bu.hao* [53.214] "not good." It takes on a rising tone [35] in (a) via the *Yi-bu-qi-ba* rule, which does not extend to case (b) where the homophonous *bu* represents a different morpheme altogether, with the meaning "division" (of an army).

[24] I owe this minimal pair to Kratochvil (1987:256).

Table 1.2. *Tonal coarticulation in Beijing*

fu-		**-ji**							
		T1	259–258	T2	209–262	T3	211–153	T4	291–162
T1	266–266	258–270 # 269–276		274–291 # 209–235		273–290 # 219–144		268–286 # 296–176	
T2	211–253	219–**245** # 266–273		216–257 # 217–249		223–281 # **247**–144		216–**243** # 281–177	
T3	214–154	213–176 # 263–272		223–168 # **188**–258		(= T2 + T3)		225–178 # 284–175	
T4	287–159	299–**218** # 262–266		300–**227** # 207–252		310–**238** # 201–141		300–**227** # 278–174	

Hyphenated numbers indicate the pitch value (in Hz) at the two endpoints of a tone shape
separates two adjacent tones
Bold letters mark coarticulation effects highlighted in Shih (1988)

4.2 *Tonal coarticulation*

In the brief sketch given above, Beijing exemplifies what has been thought of as tone sandhi in its canonical form: contextually determined tonal alternation in its simplest and purest form. However, when one examines the actual pitch curves of connected speech with the aid of laboratory instruments, one soon realizes that the tone sandhi rules formulated above produce at best an intermediate level of linguistic abstraction, which stands at a considerable distance from the pitch envelope in actual phonetic records. The full specification of pitch curve in natural, fluent speech entails at least two other types of processes. First, **tonal coarticulation**. I will limit myself to only those aspects of tonal coarticulation highlighted in Shih (1987, 1988, 1991).[25] To factor out segmental influence on tones, she constructed disyllabic expressions [*fu* + *ji*], that encode meaning contrasts exclusively in tonal terms. Thus [*fu.ji* T1.T1] means "to hatch chicks," while [*fu.ji* T1.T2] signifies "husband's residence," and so forth. Given the four-tone system of Beijing Mandarin, there are 16 such disyllabic sequences. The acoustic measurements of the pitch values of these disyllabic strings are summarized in table 1.2.

The four rows of the leftmost column correspond to the four tones of the first syllable [*fu*], pronounced in isolation; the two hyphenated numbers represent the average pitch value (in Hz) at the two end points of each tone shape. Likewise, the isolation tone values of the second syllable [*ji*] are given across the top. The phonetic shapes of the two tones in connected speech are specified in the cells where the relevant rows and columns intersect. "#" separates the two adjacent tones. Recall that

[25] Other recent experimental works on tonal coarticulation in Beijing Mandarin include Wu (1982, 1985), X. Shen (1990a, b, 1992).

Table 1.3. *Tonal coarticulation in Beijing*

fu-		**-ji**							
		T1	HH	T2	MH	T3	ML	T4	HL
T1	HH	HH # HH		HH # MH		HH # ML		HH # HL	
T2	MH	M**H**$^{-}$ # HH		MH # MH		MH # **M**$^{+}$L		M**H**$^{-}$ # HL	
T3	ML	ML # HH		ML # **M**$^{-}$H		(= T2 + T3)		ML # HL	
T4	HL	H**M** **#** HH		H**M** # MH		H**M** # ML		H**M** # HL	

T3 + T3 → T2 + T3 via the T3 Sandhi rule, hence the shaded cell. Shih (1988) used bold numbers to highlight some of the most salient and consistent tonal coarticulation effects. To bring out more clearly the essential points, I have translated table 1.2 into table 1.3, where gradient pitch values are reduced to three idealized discrete targets H(igh), M(id), and L(ow), with superscript + and – representing up- or down-shift. Thus HH, ML, and MH^{-} stand, respectively, for a high level, a low-falling, and a rising tone that undershoots its H target.

Shih singled out the following points for special comment. (i) HL falls to the baseline of the pitch range only in the final position; elsewhere it falls only to the mid-range. (ii) The low-falling ML assimilates to the preceding rising MH. (iii) The low ML depresses and delays the rise of a following tone. (iv) The rising MH does not quite reach its H target when it is following by a Hx, which abbreviates {HH and HL}, namely a tone shape with H as its first element. These tonal coarticulation rules can be stated informally as follows:

(27) i. T4 + any T — HL. T → HM

ii. T2 + T3 — MH.ML → M^{+}L

iii. T3 + T2 — ML.MH → M^{-}H

iv. T2 + {T1, T4} — MH.Hx → MH^{-}

(For clarity and simplicity, I rotate the rewrite arrow "→" by 90 degrees; the target of tonal modification is underlined. Accordingly, rule (i) is equivalent to: HL → HM / ___ T.) As we can see immediately, (i–iii) are all assimilatory in nature, and all have the effect of smoothing the transition between tonal targets. The sharp fall HL is reduced to HM by (i), since whatever tone follows, its initial target is either H or M. Rules (ii) and (iii) raise or lower the M portion of the second tone, depending on whether the first tone ends in H or L. Rule (iv) may appear to be dissimilatory in nature, but lends itself to an alternative interpretation, as suggested by Shih (1988:6f.). What is going on in (iv) is that the H target of T2 is deleted when followed by another H target. The phonetic H^{-} actually represents the transition tone between M and H. To put it differently in the more familiar format:

(28)	*fu.ji*	"medication taken orally"[26]
	MH.HL	base form
	M_. HL	Absorption
	MH^{-}. HL	Tone Interpolation

Tone Absorption, a process widely attested in African tone languages,[27] fuses two adjacent like tones. Tone Interpolation inserts a transitional pitch H^{-}, intermediate between M and H.

X. Shen (1992) proposes various diagnostics to distinguish tonal coarticulation from tone sandhi including: (a) only assimilation is considered coarticulation, but tone sandhi may be both assimilatory and dissimilatory; (b) tonal coarticulation obeys only language-independent biomechanical constraints, while tone sandhi may be subject to language-specific morphological and phonological conditions; (c) tone sandhi may effect tonemic change, while tonal coarticulation involves only allotonic variations.

It is not clear whether it is desirable or even possible to segregate tonal coarticulation from tone sandhi proper. For one thing, it is not clear that late phonetic coarticulation is exclusively assimilatory. Shih (1987:10) observes that in Standard Mandarin, the high rising [35] and falling [53] tones are phonetically *higher* when preceding a low [214] tone, apparently a dissimilatory coarticulation effect.

[26] As opposed to externally applied ointment or the like.
[27] For instance, Bamileke, Mende, Kikuyu, Hausa, Ngizim, among others, cited in Hyman and Schuh (1974).

As for Shen's second diagnostic, suffice it to say that certain late phonetic tonal perturbations may be grammatically controlled. To cite one example, Wu (1985:79f.) produces acoustic data to show that the pitch curve associated with a rising tone MH in medial position actually turns into a falling contour HM when it is followed by a low or mid tone, as illustrated by the following example

(29)	[*zhan lan*] *guan*	"exhibition hall"
	L L L	
	MH.MH.L	T3 Sandhi
	MH.HM.L	Tonal coarticulation

This fairly subtle contextual tonal perturbation has eluded most students of Chinese tonology, and is detectable only instrumentally, thanks to Wu (1982). Remarkably, such a "low-level" phonetic rule is sensitive to morphosyntactic structure: it seems to be excluded from right-branching constructions like

(30)	*he* [*leng shui*]	"drink cold water"
	H L L	base tone
	H MH.L	T3 Sandhi
	–	Tonal coarticulation

Furthermore, as we will see immediately below, what have been usually regarded as late phonetic intonational effects are also grammatically controlled.

Finally, tone sandhi literature typically covers allotonic as well as morphotonemic alternations. Take the prototypical **T3 Sandhi** in Beijing discussed above:

(31) Beijing T3 Sandhi
T3 → T2 / ___ T3
T2 = MH
T3 = L

The conventional view is that T3 merges with T2 in the specified sandhi context. Thus perceptual tests conducted by Wang-Li (1967) established that *qima* "at least" from underlying T3 + T3 becomes indistinguishable from *qi ma* "to ride a horse" from T2 + T3; likewise *fenchang* T2 + T3 "graveyard" and *fenchang* T3 + T3 "flour factory" are homophonous in actual speech. This established view has been challenged by Zee (1980) and Kratochvil (1987). Zee (p. 101) shows that underlyingly T3 preceding

another T3 has an overall lower Fo (fundamental frequency) than the corresponding T2 in the same environment. This would seem to suggest that T3 Sandhi does not entail a paradigmatic substitution of T3 by T2, but rather only turns T3 into some allotone, say T3′, in the sandhi context, and hence maintains the categorial distinction between T2 and T3′. Whatever difference may persist between T2 and T3′, it is not clear whether such difference is reliably perceived by a native speaker (Wang-Li 1967). In any event, T2 and T3′ are functionally equivalent with respect to other sandhi processes. Recall that Beijing has another sandhi rule, which I have referred to as T2 Sandhi. It turns a non-final [35] into [55], when preceded by [55] or [35]. What is important here is that this T2 Sandhi rule treats underlying [35] (T2) and derived [35] (T3′, via T3 Sandhi) alike. This is demonstrated below:

(32)	"observatory"		"watershed"	
a.	*tian wen tai*	b.	fen shui ling	
	H.MH.MH		H. L. L	base tones
	– – –		H. MH.L	T3 Sandhi
	H. H. MH		H. H. L	T2 Sandhi

Note that both of the medial syllables *wen* (underlyingly /MH/ or T2) and *shui* (originally T3 or /L/) undergo T2 Sandhi.

I will therefore take the position that there is no essential difference between tone sandhi and tonal coarticulation, except that tone sandhi processes are perceptible to the (trained but) unaided ears,[28] and therefore more likely to be reported by fieldworkers and integrated to a greater extent into the phonological component of the grammar. In practical terms, the sandhi phenomena we are about to investigate are circumscribed only by their accessibility: with the notable exception of Standard Mandarin and a small handful of better known dialects, published reports provide only more or less standardized phonetic transcriptions without

[28] Clements (1979:549) conjectures that the minimal perceptible interval is in the order of one semitone on the diatonic scale; his hypothesis is confirmed by experimental evidence: Gaoba Dong, a Kam-Tai language spoken in Guizhou, China, distinguishes five level tones (in addition to three rising and one falling tone), ranging from 129 to 259 Hz, covering roughly an octave (C_2 to C_3). Interestingly, the interval between the two lowest tones is 9 Hz, or about one semitone (see Shi et al. 1987).

In terms of discrimination between tone shapes, Dreher and Lee (1968) report that a minimum pitch change (upwards or downwards) of 2 semitones is necessary for Mandarin speakers to distinguish a rising or falling tone from a level tone. Comparable figures are 2.5 semitones for French speakers (Rossi 1971) and between 1.5 and 3 semitones in the case of Dutch subjects ('t Hart 1974). The just noticeable difference can be considerably smaller under controlled conditions, cf. Klatt (1974), Gandour (1978).

supporting acoustic data.[29] As a consequence, for all practical purposes, the "low-level" phonetic coarticulation effects of the sort illustrated above are beyond the scope of our investigation.

4.3 Intonational effects: declination and catathesis

Apart from localized coarticulation effects, tone also interacts more globally with the overall intonation pattern of the utterance. Despite its intuitive and poetic appeal, Y-R. Chao's (1968:39) metaphor of the intonational overlay upon lexical tones as "small ripples riding on large waves" is too impressionistic to make testable predictions. Since then, however, acoustic studies on tone and intonation in Chinese, basically limited to Standard Mandarin, have multiplied (Tseng 1981, Wu 1982, 1985, Gårding 1984, 1987, J. Shen 1985, Shih 1987, 1988, 1991, X. Shen 1990a, b, 1992, Liao 1994). I will only mention two types of intonational effect on tone – declination and catathesis – again liberally drawing on data from Shih.

Declination refers to the gradual downtrend of pitch over the course of an intonational phrase that is "blind to the phonological sequence of accents or tones" (Pierrehumbert and Beckman 1988:11), claimed by some to occur universally in natural fluent speech (cf. Ohala 1978, Ladd 1984, Levelt 1989). **Catathesis** or (automatic) downstep, on the other hand, is Fo lowering due to specific tonal combinations, typically the lowering effect of a L on a subsequent H in a sequence of interspersed Hs and Ls (Liberman and Pierrehumbert 1984).[30] Catathesis is attested in English (Pierrehumbert 1980), Japanese (Pierrehumbert and Beckman 1988), and Chinese. The effects of both declination and catathesis are demonstrated in an experiment that is as simple as it is elegant. Each of the four pentasyllabic sentences below has a high tone H in the odd positions, interspersed with varying tones in the even-numbered positions (Shih 1987, 1988).

[29] Notable exceptions include Shanghai (Zee and Maddieson 1980), Fuzhou (Chan 1985), Xiamen/Taiwanese (Du 1988), Wuxi (Chan and Ren 1989), and Zhenhai (Rose 1990). I have made original recordings of connected speech in the following dialects: Tianjin (north Mandarin), Pingyao (Jin), New Chongming (northern Wu), and Wenzhou (southern Wu).

[30] Catathesis or automatic downstep gives rise to phonemic downstep when the conditioning L disappears. Thus HLH $\rightarrow$ H!H, where downstepped !H contrasts with the plain H. Since Welmers (1959) drew attention to this widespread tonal phenomenon in African languages, it has been studied extensively. See Clements (1979), Hyman (1979, 1986, 1993), Clements and Ford (1981), and references cited therein.

Table 1.4.

context	average Fo of		
	H_1	H_3	H_5
Sentence a	284	265	252
Sentence b	286	255	239
Sentence c	299	260	236
Sentence d	280	259	241

Sentences a, b, c, d refer to (33)
$H_{1, 3, 5}$ refer to the high tones in the first, third, and fifth position of the sentence

(33) a. H_1.**H**.H_3.**H**.H_5
ji shi xiu tuo che "the mechanic fixes the cart"

b. H_1.**MH**.H_3.**MH**.H_5
gong ren shou fang zu "the worker collects the rent"

c. H_1.**ML**.H_3.**ML**.H_5
jing li he guo zhi "the manager drinks juice"

d. H_1.**HL**.H_3.**HL**.H_5
shang dian chu jiu shu "the store old books"

By measuring the average pitch values of the three odd-numbered H tones in each sentence,[31] we can detect the impact of both declination and catathesis, if any occurs. The results are shown in table 1.4. Sentence (a) consists exclusively of a string of Hs. Nevertheless the third and the fifth Hs (= H_3 and H_5) are pronounced on a gradually declining pitch, suggesting an appreciable declination effect.[32] An intervening low (or non-high) target tends to induce a steeper decline on the following Hs. This downstep effect is particularly robust in the case of sentence (c), where the Hs are interrupted by a low-falling ML. Thus the pitch declines by an average of 39 Hz from H_1 to H_3, and by 24 Hz from H_3 to H_5, in comparison with context (a), where the downdrifts at the same intervals are on the order of 19 Hz and 13 Hz respectively.

[31] Each sentence was recorded in random order, four times in natural speech and four times in "reiterant speech." Table 1.4 represents the averages of natural speech tokens.

[32] Curiously, neither Tseng (1981) nor Liao (1994) found a consistent declination effect in their samples consisting of mostly naturally occurring speech. Shih (1987, 1988, 1991), on the other hand, drew her conclusions from sentences specifically designed to test various hypotheses. I suspect that the sample control may have something to do with the discrepancies in their findings.

In a subsequent paper, Shih (1991) further demonstrated that declination is not a purely physiological phenomenon, but is controlled at least in part by the grammar. Specifically, she showed that declination is neither gradual nor constant, but steeper and more pronounced at word boundaries. In other words, Fo declines more gently within words, and more precipitously between words.

We are limited in our access to the relevant phonetic data on the tone–intonation interplay outside of Beijing Mandarin. A serious study of tonal behavior that integrates the results of intonational studies across dialects will have to await another occasion.

4.4 Tone sandhi vs. tone change

One final class of processes affecting tones in connected speech calls for some clarification. In all the cases we have examined so far, the context triggering tonal modification is phonological in nature, whether local (a neighboring tone, segmental influence) or global (intonation). Sometimes what conditions tone change is morphological in nature, as exemplified below:

(34) a. Taishan

ngwoi	33	"I"
ngwoi	22	"we"

b. Zhongshan

hy	22	"go"
hy	35	"gone perfective"

c. Wenling

huɔ̃	31	"yellow"
huɔ̃	15	"yolk"
baq	1q	"white"
baq	51	"albumen, egg white"

Clearly, morphotonemic alternation functions as a "process morpheme," inflectional in the case of Taishan and Zhongshan (both of the Yue group), and derivational in the case of Wenling (Wu, from R. Li 1979). Chan (1989), to whom I owe the Yue examples, observes that marking verbal aspect by tonal means is a productive process in Zhongshan. Similar morphotonemic alternations have been documented for Bobai (also Yue; Kam 1980) and other dialects groups as well, including Xiang (Tang 1960) and Min (Ting 1983), among others.[33] Sinologists generally use the

[33] For similar processes, now no longer productive, of derivation by tone change in Thai (Siamese), see Kam (1980) and work by Prapin Manomaivibool (1976) cited there.

term **tone change** (*bianyin*) to refer to morphologically conditioned tonal modification – to distinguish it from phonologically conditioned **tone sandhi** (*biandiao*). One can think of tone change thus defined as an analog of ablaut and umlaut in English functioning as both inflectional (*foot* ~ *feet*, *sing* ~ *sang* ~ *sung*) and derivational devices (*food* ~ *feed*).

The best known and most fully documented cases of tone change are Beijing Mandarin and Cantonese. A brief sketch of these two case studies will further elucidate the form and function of this phenomenon. I will suggest that one should recognize various types of "tone change," some of which are reducible to generic tone sandhi, while others are not.

4.4.1 Tone change in Beijing Mandarin

The first type of tone change is exemplified by Beijing Mandarin, which has inherited from ancient Chinese a fair number of pairs of cognate words that are related in meaning and differ only tonally. These cognates[34] reflect a once productive morphological process. Downer (1959) gathered over 200 such cognate sets from *Jingdian Shiwen*, a seventh century AD collection of readings. In each case, the etymological base carried MC tone I, II, or IV, and the derived form tone III, corresponding to modern Beijing T4 [51]. The derived forms fall into various grammatical/semantic categories, such as noun, verb, causative etc. In other words, tone III marks a deverbal noun, a denominal verb, or a causative verb derived from an adjective, and so forth. Some canonical examples in their present-day pronunciation are listed below for illustrative purposes:

(35)	a.	*zhong*	55	"center"
			51	"to hit the center of a target"
	b.	*zhong*	214	"seed"
			51	"to plant"
	c.	*yin*	55	"shade"
			51	"to shelter"
	d.	*hao*	214	"good"
			51	"to like"
	e.	*heng*	35	"horizontal"
			51	"cross-grained, hard to deal with"

Each of these doublets also constitutes a homographic pair, i.e. two lexical roots sharing one Chinese character. Others have diverged orthographically, as in

[34] Sometimes referred to as "word families."

(36)	a.	*jing*	55	"to pass through"
			51	"path"
	b.	*ming*	35	"name"
			51	"to name"
	c.	*mai*	214	"to buy"
			51	"to sell"

Yet others have parted ways phonologically, as in

(37)	*[tš^{h}uan]*	35	"to transmit"
	[tšuan]	51	"record, biography"

where aspiration was once tonally conditioned: voiced stops and affricates became voiceless aspirates in syllables taking MC tone I (corresponding to [55, 35] in Beijing); elsewhere they turned into plain voiceless stops or affricates.

Clearly, it would be inappropriate to apply the term "tone sandhi" to describe the phenomenon sketched above. It would be misleading to treat the alternation between [*hao* 214] "good" and [*hao* 51] "to like" as anything but vestigial relics of a morphological process that has long faded into the dim past.[35] Sinologists typically refer to such morphotonemic alternations by the term "derivation by tone change," apart from tone sandhi. It behooves us to maintain this conceptual and terminological distinction.

4.4.2. Tonal morpheme in Cantonese

More problematic is a second type of tone change instantiated by Cantonese that has received considerable attention in tonological literature (Chao 1947, Whitaker 1955–56, A. Hashimoto 1972:93–100, 180–187, Kam 1977, 1980, Yip 1980:60–65, Wong 1982, Bao 1990a:182–193). Recall that Cantonese has a nine-tone system (see (17)), which I reproduce below, translating the numerical notation [53, 21] etc. into the more familiar symbols [HM, ML] and so forth, for reasons that will become immediately obvious.

[35] Haudricourt (1954b) speculated that Archaic Chinese had an -s suffix, with various functions, including "transitive, causative" and the like. On this hypothesis, the [hao 214 ~ 51] alternation would have evolved in the following way:

Archaic Chinese *xâu "good" → MC xâu II = Beijing [hao 214] "good"
Archaic Chinese *xâu-s "good + transitive/causative" → MC xâu III = Beijing [hao 51] "to like"
Here II and III stand, as usual, for rising and departing tones respectively.

(38) Cantonese tone system

	level		rising	falling
	CVN	CVq		
high (*yin*)	M	Hq; Mq	MH	HM (~ H)
low (*yang*)	L	Lq	LM	ML (~ M)

To oversimplify matters considerably, the most productive type of tone change in Cantonese is one that turns a tone into a high-rising [MH] or level [H] tone in certain contexts:

(39) a. {HM, Hq} → H
b. {M, Mq, L, Lq, ML, MH, LM} → MH

I will single out three of the common contexts, illustrated with examples drawn chiefly from Hashimoto (1972). First, after the "vocative" prefixes [*a* M] and [*lou* LM]:[36]

(40) a.	*a tsæng*	M.HM	→	M.H	"old Zhang"
b.	*a ts'an*	M.ML		M.MH	"old Chen"
c.	*a tsiu*	M.L		M.MH	"old Zhao"
d.	*lou lei*	LM.LM		LM.MH	"old Li"
e.	*lou* $k^{w}ɔ{:}k$	LM.Mq		LM.MHq	"old Guo"
f.	*lou mak*	LM.Lq		LM.MHq	"old Mo"

Second, in the second syllable of a reduplicated adjective, with or without the marker *-tei* MH. Adjective reduplication has an intensification effect: AA = very A, while AA-*tei* translates as "a little A." Relevant examples follow:

(41) a.	*hung hung*	ML.ML	→	ML.MH	"very red"
b.	*yit yit*	Lq.Lq		Lq.MHq	"very hot"
c.	*p'a p'a tei*	M.M.MH		M.MH.MH	"a little scared"
d.	*lang lang tei*	LM.LM.MH		LM.MH.MH	"a little cold"

Finally, certain lexically marked items undergo tone change with concomitant shift in meaning:

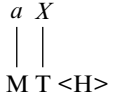
[36] Both with the connotation of familiarity and informality, roughly like "old" as in "Old Joe." Tone values are regularized to conform with *Hanyu Fangyin Zihui* (2nd edition, 1989).

(42)	a.	*t'ɔng*	ML	"sugar"	→	MH	"candy"
	b.	*nøy*	LM	"female"		MH	"daughter, girl"
	c.	*min*	M	"face"		MH	"face" (metaphorical)[37]

These morphotonemic alternations cannot be attributed to any apparent phonologically definable contextual influence. The question arises as to whether or not the term tone sandhi should be extended to encompass this second type of tone change. Pulleyblank (1991:448f.) advocates a narrow construal of the term sandhi, and excludes alternations exemplified in (40–42) from the scope of tone sandhi, which is more appropriately – on etymological grounds – restricted to the liminal changes arising out of the juxtaposition of tones.

Unfortunately, the dichotomy between morphologically and phonologically conditioned alternations is not as clear cut as one might wish. From a diachronic perspective, both types of processes often spring from the same source. Take English *foot* ~ *feet* and *food* ~ *feed*. Synchronically speaking, the vowel mutation is obviously not phonologically conditioned, but merely encodes the associated morphological information. From a historical perspective, however, umlaut (unlike ablaut) was presumably triggered by the vowel quality of the inflectional/derivational endings that have since disappeared. Thus, *feet* had its origin in pre-historic Old English **fo:t-iz* (cf. Pyles and Algeo 1982:114f.). From this angle, the morphological process of vowel mutation is not substantially different from the canonical phonological process of vowel harmony or anticipatory assimilation.

To return to the Cantonese case, one could argue, after Yip (1980:60f.), that what appears to be a morphologically conditioned tone change can be brought in line with the canonical form of phonological rules, if we make allowance for the theoretical device of floating tones. For concreteness, take the "familiar vocative" case. This construction has the following configuration:

(43)


where [a] is the M-toned vocative prefix; X, together with its lexically assigned tone T, is a place holder for the proper name, and the floating <H> represents a purely tonal morpheme – in Y-R. Chao's preautosegmental

[37] As in the surface or side of something, or as in "to lose face."

and remarkably prescient words, a "non-syllabic and non-segmental suffix" (Chao 1956:1). [*a tsæng*] "old Zhang" and [*a ts'an*] "old Chen" would then have the following as input to the phonological component:

(44) a. *a tsæng*
M HM <H>

b. *a ts^{h}an*
M ML <H>

Docking of the floating <H> on *tsæng* and *ts^{h}an*, each of which carries its lexically assigned [HM] and [ML] respectively, results in complex tones, thereby triggering a tone-simplification process. Cantonese prohibits complex tones (i.e. falling-rising or rising-falling). When complex tones arise from the linking of floating tone, they are simplified by "smoothing," namely by eliminating a medial tone segment, while preserving the two endpoints. In other words:

(45) a. *a tsæng* → *a tsæng*
M. HM<H> M. HH

b. *a ts^{h}an* → *a ts^{h}an*
M ML<H> M MH

which are exactly the pronunciations we want ([HH] on *tsoeng* is non-distinct from [H]). As for *a tsiu* "old Zhao" and *a lei* "old Li," linking the floating <H> to the proper names would create [LH] and [LMH], the latter simplifying to [LH].

(46) a. *a tsiu*
M. L<H>

b. *a lei* → *a lei*
M. LM<H> M LH

At this point several logical moves are open to us. The most theory-neutral option is to say that Cantonese distinguishes only two rising

tones: high [MH] or low [LM]; as a consequence, the *tertium quid* [LH] is reduced to one or the other permissible contours. As it turns out, Cantonese opts for [MH] as the alternative of choice.[38]

The treatment of the "familiar vocative" can be extended, *mutatis mutandis*, to cover adjective reduplication and lexically marked tone changes. Notice, however, that this account of the "familiar vocative" makes crucial use of an underlyingly floating <H>. There is, for instance, no plausible dummy syllable occupying the slot after the proper name, that could undergo deletion leaving behind a floating <H>.[39] The question then turns on whether it is justifiable, synchronically speaking, to posit a {familiar vocative} morpheme that is purely tonal, i.e. without syllabic mooring, in the *underlying* representation.

The answer to this question is somewhat ambiguous. While purely tonal morphemes are certainly atypical in Chinese,[40] the legitimacy of underlyingly floating tones is unassailable on theoretical and cross-linguistic grounds. Close analogs of the {familiarity} morpheme in Cantonese are not hard to find outside Chinese. Perhaps the best known example is the associative particle in the Chadic language of **Ga'anda** as described in Kenstowicz (1994:363f.) based on Ma Newman (1971). The essential facts are given below:[41]

(47) a. *al* "bone"
M

b. *cunewa* "elephant"
L L L

c. *al cunewa* "bone of elephant"
M H L L

[38] Another option is to ban "cross-register" contours, including LH. This move is theory-dependent to the extent that it presupposes a tonal geometry along the lines proposed by Yip (1980). See chapter 2.

[39] Cantonese does have a suffix-like [-a] which is, however, an unlikely candidate as the dummy syllable. Consider for instance:

a tsoeng a → *a tsæng a* "old Zhang"
M HM M M HH M

There are two decisive reasons why this final particle -a cannot be the trigger of tone change: (i) tone change is independent of -a deletion: HM changes to H regardless of whether -a is present or absent; and most importantly, (ii) this final particle [-a] carries a M not a H tone when it does appear.

[40] Other hypothetical cases of floating tone are Standard Mandarin "vivid" adjective reduplication and Xiamen triple reduplication. See Yip (1980:66f.) for details.

[41] [b] stands for implosive [b].