

A Grammar of Mandarin Chinese

普通话语言学

By

林华

Hua Lin, Ph. D.

© Hua Lin 2001

Dr. Hua Lin

Associate Professor

Department of Linguistics

University of Victoria

P.O. Box 3045

Victoria, BC

Canada, V8W 3P4

Phone: 250-721-6643

Fax: 250-721-7423 (attn to me)

E-mail: Luahin@uvic.ca

Home Page: <http://web.uvic.ca/ling/lin/www>

Book published by: Lincom Europa, Germany

ISBN: 3895866423

The book is a linguistics introduction and study of the language called Mandarin Chinese. Its five chapters cover the history, phonetics, phonology, morphology, parts of speech and syntax of the language. In this file, you will find

Key words: 汉语语言学 普通话语言学 国语语言学

Chinese Linguistics, Chinese phonetics, Chinese phonology, Chinese morphology, Chinese syntax, Chinese dialects

This PDF file contains

Chapter 2 Phonetics and Phonology

Pages 23-51 (exact numbers of the internal pages may differ a slightly from the hard copy version)

Chapter 2. Phonetics and Phonology

The study of the sound system of language normally starts with phonetics followed by phonology. However, it is widely acknowledged that the two are not mutually exclusive: The quality of phonological modeling crucially depends on the understanding of phonetic details, whereas phonetic description is often more insightful when done within a more explanatory framework of phonological constraints, rules and/or parameters. In this chapter, we will not attempt to separate the two, but each will be resorted to as required by the description and explanation of certain aspects of the Mandarin sounds and their patterns.

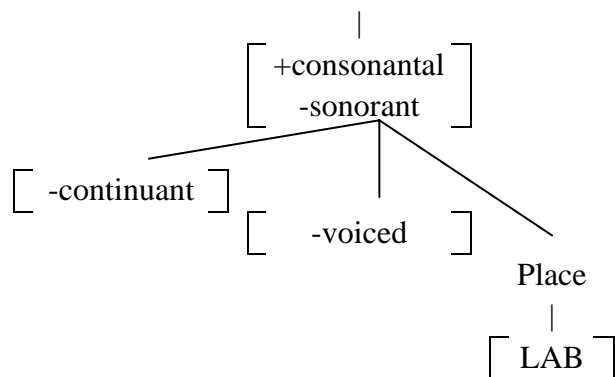
The theoretical framework used in this chapter will be Generative Phonology, a major influential phonological framework developed in the twentieth century. Within this framework, a sound (or in proper linguistic terms, a phone or a segment) is a bundle of binary features which encode such information as the position where the sound is produced (where the stricture is) by the speech organs and the manner in which the sound is uttered. Phonological processes are captured in terms of change in the configuration of the feature bundles in certain contexts. This theory of the sound patterns of language has gone through a great deal of change within the past two decades. An example of such change is that the place features, which used to be binary in standard Generative Phonology (Chomsky and Halle 1968), are now unitary features. The arrangement of features has also been completely revolutionized. To denote a segment, say [p], the features used to be placed, unordered, in a two-dimensional feature matrix such as this,

$$(2.1) \quad \left[\begin{array}{l} +\text{consonantal} \\ -\text{sonorant} \\ -\text{continuant} \\ -\text{voiced} \\ +\text{labial} \\ \dots \end{array} \right]$$

but they are now organized into a three-dimensional tree structure which branches out into several levels from a root node carrying the crucial information as to whether the segment is a consonant or a vowel. The same information given in (2.1) would be organized roughly as (2.2):¹

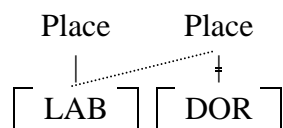
¹ This is a much simplified feature geometry tree, though adequate to illustrate here the difference between the old and new treatment of features. Interested readers may read Broe (1992) and Halle (1992) for details.

(2.2)



Phonological processes used to be described by rules in the form of 'A becomes B in the environment of C', but now may be indicated by lines linking among the nodes of features. The following dotted line represents a phonological process whereby the second segment assimilates the place feature of the first.

(2.3)



This newer version of Generative Phonology is more complex and abstract. The understanding of it entails a great deal of background knowledge of the various aspects of the theory. Because the scope of this descriptive grammar does not allow a full introduction of this new theory, for ease of comprehension on the part of the general reader, we will now and then adhere to more traditional ways of describing phonetics and phonological processes. The newer theory will be used where necessary, especially where the older theory is silent about the issues involved or is entirely inadequate.

This chapter addresses four aspects of the sound system of Mandarin: sounds, syllables, segmental processes and tone. The section on sounds will include the consonants, vowels, and glides. The section on the syllable will address such topics as the initials and finals, phonotactic constraints, syllable weight, etc. The third section discusses the phonological processes that occur between adjacent segments within a syllable. This section will also address the phonological versus phonetic representation of phonemes and allophones in the Pinyin transcription system. The section on tone consists of two parts, one devoted to the four basic tones and the other to the neutral tone.

2.1 The Sounds

A description of the phonetics and phonology of a language cannot be done without

resorting to the linguistic tool called the phonetic features. Each sound or segment² is identified by a unique set of distinctive features organized in a tree structure. Features can be contrastive in one language, but not in another. Some of the most important distinctive features for the description of Mandarin include 1) major features such as [\pm consonantal], [\pm vocalic], [\pm aspiration], and [\pm nasal]; 2) place features: [LABIAL], [CORONAL], and [DORSAL]; and 3) vowel features: [\pm high], [\pm back], and [\pm round].³

2.1.1 Consonants

The phonetic features that distinguish consonants from vowels are [\pm consonantal] and [\pm vocalic]. While consonants are [+consonantal, -vocalic], vowels are exactly the opposite: [-consonantal, +vocalic]. There is a third group of phonemes called glides or semi-vowels that have negative values for both features: [-consonantal, -vocalic]. Strictly speaking, Mandarin has 20 phonemic consonants.⁴ Traditionally, the [-consonantal] glides are grouped with the consonants. As the glides [w, y, ʅ] are not phonemic in Mandarin, we have chosen not to include them here, but will deal with them at a later point in this chapter.

(2.4)

<i>Manner</i> \ <i>place</i>		<i>Labial</i>	<i>dental</i>	<i>alveolar</i>	<i>alveo-palatal</i>	<i>palatal</i>	<i>velar</i>
		<i>Stop</i>	[+asp]	pʈ		tʈ	
	[-asp]	p		t			k
<i>affricate</i>	[+asp]		tsʈ		tʃʈ		
	[-asp]		ts		tʃ		
<i>fricative</i>	[+asp]	f	s		ʃ		x
	[-asp]				ʃ̥		
<i>approximant</i>	nasal	m		n			ŋ
	liquid			l			
	retroflex			ɭ			

Two major parameters are used in phonetics to differentiate sounds in language: place of articulation and manner of articulation. An important feature in the dimension of manner of articulation in Mandarin is [\pm aspiration] ([\pm asp] for short), denoted here by a superscript

² In autosegmental phonology, the segment does not have to be a fully specified pronounceable segment, but a partial one, a theory that will not be expounded here.

³ These vowel features are terminal features of [LAB] ([round]) and [DOR] ([high] and [back]). See Broe (1992) for details.

⁴ The three palatal obstruents [t̪ä, t̪äʈ, ä], which normally appear in a Mandarin table of consonants, are excluded here because they are not, strictly speaking, phonemic. They will appear later in a more inclusive table in the section on variations of phonemes (Section 2.3).

diacritic⁵ 'J', that divides the obstruents (stops, affricates and fricatives) into two groups: those that are aspirated and those that are not. In physical terms, aspirated consonants involve a strong puff of air coming out of the lungs, whereas unaspirated consonants do not have this trait. Regarding the difference between the alveopalatal fricatives [ʃ] and [ʃ̂], there has been controversy. Textbooks and grammar books published in the People's Republic of China (e.g., Wang and Pang 1996, p. 156) distinguish the two in terms of [voicing]; namely [ʃ] is voiceless while [ʃ̂] is voiced. However, that would mean the positing of a voicing feature which is otherwise unwarranted for the Mandarin obstruents. One way to avoid this theoretical inadequacy is to assume that [aspiration] is the feature which distinguishes the language's fricatives as with the other obstruents. This is the position held by many linguists outside China, and is adopted in the table above. It is interesting to note that [aspiration] as a distinctive feature is often hard to comprehend by speakers of a language making a distinction among its obstruents in terms of voicing rather than aspiration. For instance, to a speaker of English, the aspiration distinction in Mandarin sounds indistinguishable from the voicing distinction in English. The reverse situation is true for native-Mandarin-speaking learners of English.

From the point of view of place of articulation, the consonants can be classified into three major groups: [LABIAL] (or [LAB] for short) which include the labials, [CORONAL] (or [COR] for short) which include the dentals, the alveolars and the alveopalatals, and [DORSAL] (or [DOR]) which include the rest. The [COR] alveopalatals [tʃ, tʃ̂, ʃ, ʃ̂] are a special case. These sounds are similar to English [tʃ, dʒ, ʃ, ʒ] (the initial consonants in *church*, *judge*, *ship* and *genre*, respectively) in place and manner of articulation, but are different from their English counterparts in that they are pronounced without the protruding of the lips but instead with the additional feature of curling the tip of the tongue toward the front of the hard palate. The latter characteristic makes them [retroflex] sounds. Mandarin spoken as a second dialect in non-Mandarin-speaking areas, particularly in the Southeast provinces of China, normally does not have this [retroflex] feature. Taiwan Mandarin, for one, clearly lacks this feature.

2.1.2 Vowels

Mandarin has a relatively small set of phonemic vowels, compared with some other Chinese dialects (e.g., Shanghai). There are six phonemic vowels⁶ which are given below:

(2.5)

	<i>front</i>		<i>central</i>	<i>back</i>	
	<i>unround</i>	<i>round</i>		<i>unround</i>	<i>round</i>
<i>high</i>	i	ü			u
<i>mid</i>				ɿ	o
<i>low</i>			a		

⁵ Another very common way of transcribing aspiration is the use of the 'prime' sign, with which the aspirated [p] would be [pʰ].

⁶ The vowels after the sibilants [tsʃ, ts, s, tʃ̂, tʃ̂̂, ʃ̂, ʃ̂̂] are dealt with later in this chapter.

Three features are sufficient in distinguishing the phonemic vowels in Mandarin: [backness] (front/back), [height] (high/low) and [roundness] (round/unround). Unlike many other Chinese dialects, Mandarin has two high front vowels, [i] and [ü]. While [i] is a very common vowel similar to the English vowel in *beat*, [ü], which is pronounced with the tongue positioned as in [i] and the lips positioned as in [u], is not as common. Not only is it often absent from other dialects of Chinese, but it is also occasionally absent from Mandarin as spoken in Taiwan. Thus, one sometimes hears *nǐpéngyǒu* 你朋友 '(literally) you friend' in place of the targeted *nǚpéngyǒu* 女朋友 'girl friend' in Taiwan Mandarin.

The mid-vowel /ɨ/ is often substituted for the English schwa [ɜ] by an English-as-a-second-language (ESL) learner whose mother tongue is Mandarin. Both are non-rounded mid-vowels, but the Mandarin one is tenser and further back in the oral cavity. The low vowel /a/ is fairly similar to the English vowel [a] in *father*, the mid vowel /o/ to that in *boat*, and the back vowel /u/ to that in *boot*. All five vowels have contextual variations, a topic to which we will return shortly.

2.2 The Syllable

One of the most remarkable phonological characteristics of Mandarin (and the Chinese language in general) is the salience of its syllable. The following quote from Jerry Norman (1988) is typical of linguists of Chinese:

'There is a sense that Chinese is also phonologically monosyllabic. In almost all descriptions of Chinese, the syllable is taken as a kind of self-contained entity which forms the basis of phonological description... In historical comparison, [the syllable] is the largest relevant unit; another important feature of Chinese dialects (and perhaps of other monosyllabic languages as well) is that any one dialect contains a fixed number of possible syllables. Even when new terms are borrowed from foreign languages, they are interpreted in terms of the existing set of syllables... A further consideration is that most phonological processes affect the syllable without reference to its lower level constituents.' (p: 138)

What is also remarkable about the Mandarin syllable is that there is a very limited number of them, just a little more than 400—the official figure from the PRC is 405—when tonal difference is disregarded. These 405 syllables are given in the Table of Beijing Sounds and Syllables below:

(2.6)

Standard Table of Beijing Sounds and Syllables

Initials	Open-Mouth (A)														Close-Teeth (B)							Round-Lip (C)								Tense-Lip (D)													
	a	o	e	-i	er	ai	ei	ao	ou	an	en	ang	eng	ong	i	ia	iao	ie	iou	ian	in	iang	ing	iong	u	ua	uo	uai	uei	uan	uen	uang	ueng	ü	üe	üan	ün						
b	ba	bo				bai	bei	bao		ban	ben	bang	beng		bi		biao	bie		bian	bin		bing		bu																		
p	pa	po				pai	pei	pao	pou	pan	pen	pang	peng		pi		piao	pie		pian	pin		ping		pu																		
m	ma	mo				mai	mei	mao	mou	man	men	mang	meng		mi		miao	mie	miu	mian	min		ming		mu																		
f	fa	fo					fei		fou	fan	fen	fang	feng												fu																		
d	da		de			dai	dei	dao	dou	dan		dang	deng	dong	di		diao	die	diu	dian			ding		du		duo		dui	duan	dun												
t	ta		te			tai		tao	tou	tan		tang	teng	tong	ti		tiao	tie		tian			ting		tu		tuo		tui	tuan	tun												
n	na		ne			nai	nei	nao	nou	nan	nen	nang	neng	nong	ni		niao	nie	niu	nian	nin	niang	ning		nu		nuo		nuan							nü	nüe						
l	la		le			lai	lei	lao	lou	lan		lang	leng	long	li	lia	liao	lie	liu	lian	lin	liang	ling		lu		luo		luan	lun					lü	lüe							
z	za		ze	zi		zai	zei	zao	zou	zan	zen	zang	zeng	zong											zu		zuo		zui	zuan	zun												
c	ca		ce	ci		cai		cao	cou	can	cen	cang	ceng	cong											cu		cuo		cui	cuan	cun												
s	sa		se	si		sai		sao	sou	san	sen	sang	seng	song											su		suo		sui	suan	sun												
zh	zha		zhe	zhi		zhai	zhei	zhao	zhou	zhan	zhen	zhang	zheng	zhong											zhu	zhua	zhuo	zhuai	zhui	zhuan	zhun	zhuang											
ch	cha		che	chi		chai		chao	chou	chan	chen	chang	cheng	chong											chu	chua	chuo	chuai	chui	chuan	chun	chuang											
sh	sha		she	shi		shai	shei	shao	shou	shan	shen	shang	sheng												shu	shua	shuo	shuai	shui	shuan	shun	shuang											
r			re	ri				rao	rou	ran	ren	rang	reng	rong											ru		ruo		rui	ruan	run												
j															ji	jia	jiao	jie	jiu	jian	jin	jiang	jing	jiong														ju	jue	juan	jun		
q															qi	qia	qiao	qie	qiu	qian	qin	qiang	qing	qiong															qu	que	quan	qun	
x															xi	xia	xiao	xie	xiu	xian	xin	xiang	xing	xiong																xu	xue	xuan	xun
g	ga		ge			gai	gei	gao	gou	gan	gen	gang	geng	gong												gu	gua	guo	guai	gui	guan	gun	guang										
k	ka		ke			kai	kei	kao	kou	kan	ken	kang	keng	kong												ku	kua	kuo	kuai	kui	kuan	kun	kuang										
h	ha		he			hai	hei	hao	hou	han	hen	hang	heng	hong												hu	hua	huo	huai	hui	huan	hun	huang										
	a	o	e		er	ai	ei	ao	ou	an	en	ang	eng		yi	ya	yao	ye	you	yan	yin	yang	ying	yong	wu	wa	wo	wai	wei	wan	wen	wang	weng	yu	yue	yuan	yun						

Even when tonal difference is considered as a difference in syllable type, there are only about 1200 types. This number is remarkably small in contrast to that in English, which has more than 8,000 syllable types according to one source (De Francis 1984). One consequence of this paucity of syllable types is that in Mandarin there are a great many homophones, especially at the morpheme level. We will return to this topic of homophones in the next chapter.

2.2.1 The Initial And the Final

The Chinese syllable has been the substance of scholarly studies for centuries in China. As particularly evident in such rhyme dictionaries as *Qieyun* (Lu Fayan, A. D. 601), a Chinese syllable is traditionally seen as having two parts: the initial (*shēngmǔ* 声母) and final (*yùnmǔ* 韵母).⁷ The initial is usually the single consonant found at the beginning of a syllable while the final everything that follows. The initial may be empty, or a 'zero initial' (*língshēngmǔ* 零声母) in which case, the initial consonant is absent, and the syllable begins with a vowel⁸ which is still considered to be part of the final.

The division of a syllable into just two parts seems to be psychologically real in the subconscious linguistic knowledge of native Mandarin speakers. Studies (e.g., S. Wang 1993) have shown that Chinese who have not learned an alphabetical system such as Pinyin or the writing system of English cannot further segment a Chinese syllable final into its component sounds⁹—suggesting the integrity of the Chinese final. This bipartitional approach to syllable segmentation is still largely used in the dialect studies of Chinese in the People's Republic of China. There, in almost all books of dialect studies—a great number of which have been produced during the last ten years, one can find lists of all the initials and all the finals of the dialects under study.

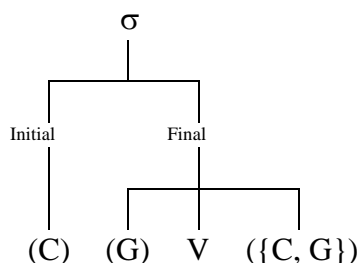
Scholars of the Chinese language perhaps would have been content with the initial-final partition of the Chinese syllable, had not Western linguistics been brought to China by such pioneering modern linguistic researchers as Yuen Ren Chao (or in Pinyin, Yuanren Zhao) early in the Twentieth Century. Today, many grammar books published in the People's Republic of China reflect the influence of modern linguistics from the West. One example is that in many grammar books, the Mandarin syllable is described in finer details, which we have distilled in the following notation:

⁷ *Qieyun* 切韵 also indicates the existence of tone over the syllable. We will address tone later in this chapter.

⁸ The vowel may change into a glide. See the sections on vowels.

⁹ Samuel Wang found that some native speakers of Taiwanese, a subdialect of Min (or Fukienese) could not even segment a Taiwanese syllable into its initial and final. However, given the existence of many secret languages, language games and speech errors which move finals around, it is indisputable that initial-final segmentation is psychologically real.

(2.6)



Where G is a contextual variation of a high vowel, and C in the Final is a nasal consonant.

Among the symbols, 'σ' represents a syllable, C a consonant, V a vowel, and G a glide (which is a contextual variation of a high vowel). The use of the braces indicates that all but the nucleus vowel V is optional in a syllable; the use of the curly brackets means that only one of the members enclosed can appear at a time. The above schema, if spelt out, yields the following 12 syllable structures in Mandarin. For consistency, all examples given are first-toned monosyllabic words:¹⁰

(2.7)

<i>Syllable Types</i>	<i>Examples</i>		
	<i>IPA</i>	<i>Pinyin</i>	
<i>CGVC</i>	pyGn	bian	边 'side'
<i>CGVG</i>	~wGy	shuai	摔 'break'
<i>CVC</i>	fHP	feng	风 'wind'
<i>CVG</i>	xey	hei	黑 'black'
<i>CGV</i>	täya	jia	家 'family'
<i>CV</i>	tJa	ta	他/她/它 'he/she/it'
<i>GVC</i>	yGn	yan	烟 'smoke'
<i>GVG</i>	wGy	wai	歪 'not straight'
<i>GV</i>	ya	ya	鸭 'duck'
<i>VC</i>	an	an	安 'peace'
<i>VG</i>	aw	ao	凹 'convex'
<i>V</i>	u	wu	屋 'house'

The three parts which form the final of a syllable have been referred to as the 'head' *yùntóu* 韵头, (or 'medial' *jièyīn* 介音), the 'middle' *yùnfǔ* 韵腹 and the 'tail' *yùnwěi* 韵尾

¹⁰ Please refer to (2.16) for the use of phonetic symbols for glides (e.g., [y]) in this chapter.

respectively. The head is a high vowel variation of a glide [y], [w] or [ÿ].¹¹ The middle is any one of the vowels when a glide is not present. When it is, the middle is a non-high vowel. The tail has four forms: a glide [y] or [w], or a nasal [n] or [ŋ].

Depending on what sound starts them, the finals have traditionally been classified into four groups: 1) the 'open-mouth finals' (*kāikǒuhū* 开口呼) which begin with a non-high vowel (i.e., [a], [o] or [ɤ]), 2) the 'close-teeth finals' (*qíchǐhū* 齐齿呼) which begin with the high front unround vowel [i], 3) the 'close-mouth (lip-rounding) finals' (*hékǒuhū* 合口呼) which begin with the high back rounded vowel [u], and 4) the 'tense-lip finals' (*cuōkǒuhū* 撮口呼) which begin with the high front rounded vowel [ü]. The columns of the Standard Table of Beijing Sounds and Syllables, given in (2.8), are normally arranged based on this four-way division. Mandarin finals have also been classified into simple finals, complex finals, and nasal finals. A simple final contains only one vowel. A complex final consists of a diphthong (two-vowel sequence¹²) or a triphthong (three-vowel sequence). A nasal final ends in a nasal consonant. In Pinyin, the Mandarin finals are given below:

(2.8)

<i>Type of Finals</i>	<i>Open-Mouth</i>	<i>Close-Teeth</i>	<i>Close-Mouth</i>	<i>Tense-Lip</i>
<i>Simple</i>		i	u	ü
	a	ia	ua	
	o		uo	
	e	ie		üe
<i>Complex</i>	ai		uai	
	ei		uei(ui)	
	ao	iao		
	ou	iou(iu)		
<i>Nasal</i>	an	ian	uan	üan
	en	in	uen(un)	ün
	ang	iang	uang	
	eng	ing	ueng	
	ong	iong		

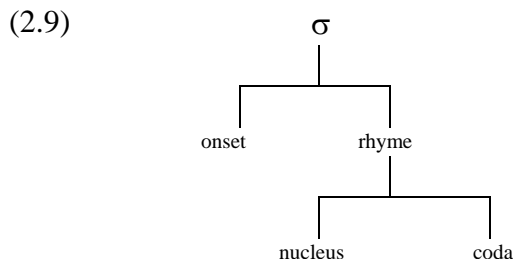
As shown above, a number of finals contain diphthongs. A noteworthy property of these two-vowel sequences is that, compared with similar strings in English, they are much more fused. A final with them often sounds more like single vowels (or monophthongs) to the ears of a native English speaker. For instance, the <ai> in 'Shanghai' is said with a lot less transition from the vowel to the glide than the similar [ay] in 'bye' in English. Thus, a native English speaker, when saying the word 'Shanghai', sounds to be exaggerating the <ai> to a native Mandarin speaker.

¹¹ These glides are sometimes represented by [j], [w], [ÿ] respectively in previous works on Mandarin phonetics.

¹² One of the vowels appears in the form of a glide (see later in this chapter for details).

Another notable matter in the above table is that there is a special final that is not included in it: <er>. The reasons for the omission are: One, it does not belong to any of the groups in the table. Two, morphemes/words that take <er> as their syllable final are so few as to be marginal. Three, it can only occur alone—neither with an initial consonant nor with a prevocalic glide.

The traditional initial-final division of the Mandarin syllable resembles the well-cited analysis of syllable structure in modern linguistics, as given below:



In this analysis, a syllable is first partitioned into an 'onset' and a 'rhyme' (also spelt 'rime'). The rhyme is further divided into a 'nucleus' vowel (or the 'peak') and a 'coda'—the final consonant or consonants of the syllable. Initially, one may think that 'onset' and 'rhyme' can be regarded as being the same as 'initial' and 'final'. However, they are not quite identifiable as will be seen in what follows.

Several questions about the constituents of the Mandarin syllable can be raised within the syllable analysis in (2.9). One of the most well-known is where the prevocalic glide goes. This question has never been raised within the traditional initial-final framework, in which the glide is clearly indicated to be part of the final. There are at least two schools of theory regarding this issue. One treats the prevocalic glide as a secondary feature of the onset (S. Duanmu 1990), the other, the more traditional, considers it part of the nucleus. To understand this controversy, we need to digress a little to see another important characteristic of Mandarin syllables: that they all weigh the same.

It is clear from the table in (2.8) that Mandarin syllables may contain from 1 to 4 sounds, and there are altogether twelve patterns for their combination. However, these facts should not be construed as meaning that Mandarin syllables vary in weight. The truth is that, unlike many other languages in which the numbers of Cs and Vs (or Gs) in a syllable matter in the weight of the syllable, the weight of the Mandarin syllable remains constant across syllable types within a specific context of utterance. Thus, the metrics in Mandarin poetry do not normally count the number of sounds in a syllable, although they require the tone (a topic which will be addressed later) to be more or less of a certain type in a given environment.

During the last decade, linguists have tried to capture the same-weight phenomenon in more formal terms. One way of doing this is to say that within the syllable, there is a fixed number of timing slots. For instance, [ˊwGy] 'break', [ya] 'duck' and [u] 'house' would all have three timing slots within the syllable:

$$(2.10) \quad \begin{array}{c} \sigma \\ /|\backslash \\ \text{X X X} \end{array}$$

Depending on the theory one adheres to, the timing of a syllable can be analyzed and represented in several ways. What is shown above is one which uses a 'place holder' symbolized by 'X', to encode one timing unit within a syllable. With just three timing units, a question emerges immediately: how should the connection be made between these timing units and the actual Cs, Vs and Gs in Mandarin? The following illustrates the problem more visually:

$$(2.11) \quad \begin{array}{ccc} \sigma & \sigma & \sigma \\ /|\backslash & /|\backslash & /|\backslash \\ \text{X X X} & \text{X X X} & \text{X X X} \\ \dots & \dots & \dots \\ \sim \text{wGy} & \text{ya} & \text{u} \end{array}$$

If the first timing slot 'X' is in the onset position, and the second and third in the rhyme position,¹³ where should the medial glide (e.g., [w] in [\sim wGy]) go? Is it part of the timing of the onset, the first slot, or part of the timing of the rhyme in the second slot? As mentioned earlier, one theory holds that the medial glide is part of the first slot. To be more specific, it is a secondary feature of the onset. Or, if the initial consonant is absent, it is the onset itself. If the medial glide is [y], then the onset has a secondary feature of palatalization ([+high, -back]), and if it is [w], the onset has a secondary feature of [+round]. For example, the onset in the syllable [pyGn] 'side' would be [p^y], and the onset in [\sim wGy] 'break' would be [\sim w]. The two syllables would be analyzed thus:

$$(2.12) \quad \begin{array}{cc} \sigma & \sigma \\ /|\backslash & /|\backslash \\ \text{X X X} & \text{X X X} \\ ||| & ||| \\ \text{p}^y \text{ Gn} & \sim^w \text{ Gy} \end{array}$$

With the medial glide taken care of in this manner, the rest of the components in the syllable all fit into the syllable structure. This theory seems to have neatly solved the problem of the position of the prevocalic glide. Yet, it has done so at a high price. Such uniform treatment of the medial glide leaves unexplained a number of observations about the language. In secret

¹³ Not all theories assume a multi-level representation of the syllable as shown in (2.9). Some assume a 'flat' structure without the intermediate level of onset/rhyme division. The difference is not crucial in our discussion, so it will be ignored here.

language, language play and speech errors in Mandarin, it is very common to see exchange or substitution of rhymes that are complete with the medial glide. In these language processes, the medial glide behaves quite obviously as a part of the rhyme. Also, to a native speaker, treating the medial glide as a mere secondary feature of the onset is quite counter-intuitive. Both the [y] and the [w] in the above two examples, for instance, require a fair amount of time to utter. They are by no means uttered simultaneously with the onset consonant, as a secondary feature is expected to be. It seems that the fault of this theory stems from a Procrustean bed or what a Chinese saying describes as 'cutting the feet to fit the shoes': trying to fit the language data into a theory rather than developing a theory based on the data. In brief, even though we know that the Mandarin syllable weighs the same across categories, how the weight is distributed across the phonemes in the syllable remains an outstanding problem.

2.2.2 Phonotactic Constraints

The twenty consonants in the table given in (2.4) do not occur in all positions within the syllable. Except [P] and [ǒ], all occur at the beginning of the syllable. The [DOR] nasal consonant [P] occurs only at the end of a syllable in Standard Mandarin, although it can be an important syllable-initial consonant in other dialects spoken in the greater Mandarin area (e.g., Jinan of Shandong Province). Another consonant that only occurs syllable-finally in Mandarin is the retroflex approximant [ǒ], which only combines with the mid-vowel [H] to yield the form [Hǒ] (Pinyin <er> mentioned previously in our discussion of Mandarin finals) for a very few lexical items; and it often appears as a sub-syllabic suffix¹⁴ with a diminutive meaning. In fact, while all vowels (and glides) appear syllable finally, there are very few consonants that can occur at the end of a syllable. In addition to [P] and [ǒ], there is the [COR] nasal [n]. This latter nasal consonant stands out from the rest of the consonants in that it is the only one that appears both at the beginning and end of a syllable (e.g., *nán* 男 'male'). Within the syllable, there is another important feature in Mandarin: there are no consonant clusters; that is, two consonants never occur adjacent to each other in a syllable. Immediately following an initial consonant, there can be a vowel which, if followed by another vowel, is a high vowel that has become a glide. Both high and low vowels can begin and end a syllable, but when there are more than two vowels in the syllable, one of them is always a high vowel. A glide which is the [-consonantal] variation of the high vowel may begin or end the syllable.¹⁵

2.2.2.1 Syllabic Consonants

We have learned that if a syllable has an initial consonant, that consonant always has to combine with a vowel to form a syllable. That is, a syllable has to have a vowel as seen in (2.6). However, there is a group of Mandarin consonants which can form the peak of a syllable and behave just like a vowel. These consonants are called syllabic consonants. Mandarin has two sets

¹⁴ Mandarin, as all Chinese dialects, has no sub-syllabic suffix except for this one. We will address this matter further in the next chapter.

¹⁵ We will explain the vowel and glide correspondence in the next section.

of such consonants: the dentals [tsʃ, ts, s] (Pinyin <c, z, s>) and the alveopalatals [tʃ, tʃ̥, ʃ, ʃ̥] (Pinyin <ch, zh, sh, r>). These consonants can form syllables on their own by being prolonged and voiced. The voiced part gives these consonants the vowel-like and syllabic quality. The representation of such syllables has always been challenging. One way of doing it is just to use the consonant symbol with or without the diacritic 'ɿ' underneath it to show that that consonant is syllabic. For instance, in Cantonese, a syllable with a nasal syllabic consonant is normally represented with just the consonant itself (e.g., [p̚] for the Cantonese name *Ng*). Another way of doing it is to use a vowel symbol after the consonant (e.g., [tsɿ], [tʃ̥ɿ], and sometimes [tsz¹⁶]). This is actually a very common practice among linguists of Chinese (e.g., C. C. Cheng 1973 and Lin and Wang 1992). They normally use the IPA symbol [ɿ] for the vowel after the dentals, and [ɿ] for the vowel after the alveopalatal retroflexes. Both vowels, according to Lin and Wang, are apical (tongue-tip) vowels. However, for the Mandarin syllables with syllabic consonants, there are four reasons that a vowel symbol should not be used.¹⁷ The first is that using a vowel symbol gives the misleading indication that the consonant onset and the following vowel are two individual, unrelated sounds, and thus misses the fact that the voiced part of the syllable is just a continuation of the previous consonant. Secondly, the approach fails to explain why the vowel cannot take any other consonants as its onset. Thirdly, it fails to explain why the vowel cannot take a coda (a final consonant), while all the other vowels can. Fourthly, it fails to explain why the vowel cannot stand alone to form a syllable while all other vowels in Mandarin can. Lastly, when liaison occurs between a syllable and a following one with [a] (See Section 2.3), the added consonant initial of [a] is clearly a derivative of the preceding consonant—sharing at least its place of articulation—if the consonant is a syllabic one. Otherwise, the initial consonants of the two syllables have nothing to do with each other. Thus, in a tree-structure analysis, these syllables should be represented as follows:

(2.13)	σ	σ	σ	σ
	/\	/\	/\	/\
	C V	C V	C V	C V
	/	/	/	/
	ts̥	tsb	tʃ̥	tʃ̥b
	'to spray'	'capital'	'to eat'	'to know'

2.2.2.2 Syllable Gaps

Not all Mandarin consonants combine with all vowels to form a syllable, as the Standard Table of Beijing Sounds and Syllables (see ()) shows. The missing combinations create two types of gaps in syllable types. One type of gap is created when a natural class of consonants does not combine with a certain final. For instance, the whole class of labial obstruents <b, p, m, f> fails to

¹⁶ This last symbol [z], of course, is not a conventional vowel symbol, but it has been used by linguists of Chinese for the purpose in discussion.

¹⁷ See Section 2.3.3 for more discussion.

combine with the 'open-mouth' final <ong>, even though all the rest of the classes of consonants do combine with it. There are also gaps in which a natural class of consonants does not combine with a group of finals beginning with a certain vowel. For instance, all labial obstruents <b, p, m, f> fail to combine with finals beginning with the high-front-rounded vowel <ü>. The other type of gaps, on the other hand, are created when a certain consonant fails to combine with certain final even though all the rest of the consonants within the same natural class do combine with that final. An example of such a syllable gap is <bou> which does not exist even though <pou, mou, fou> do. Another example is the absence of <shong>: The retroflex <sh> cannot combine with <ong> while the rest of the retroflex consonants can: <zhong, chong, rong>.¹⁸ Presumably, the second type of gaps may be more of an accident than the first, and therefore, may potentially enter the syllable set more easily.

Synchronically speaking, however, syllable gaps are seldom, if ever, filled in Mandarin; the set of Mandarin syllables (as given in the Table of Mandarin Sounds) is closed. Syllables that do not exist today are ill-formed and rejected by native speakers. The relative uniqueness of this situation is better understood if we consider the accidental gaps in, say, English. In this language, a native speaker will accept a gap such as 'blick' as being a well-formed syllable, even though it is not in use—there is no meaning attached to it. The same is not true for the gaps in Mandarin. If not in use, it is not normally accepted. The closedness of the syllable set in Mandarin is manifested in the transliteration of foreign words, as mentioned earlier in the quote from Norman (1988). Foreign syllables (and sounds) are not allowed into the system, but are normally rendered into syllables that are members of the closed set. Diachronically speaking, when a two-syllable word contracts into a single-syllable one, the output syllable is usually among the synchronically acceptable set and not the phonological combination of the two original syllables.

2.3 The Processes

One important characteristic of Mandarin (as well as other Chinese dialects) is that segmental phonological processes (i.e., processes not involving tone) almost always occur within the syllable. One rare exception is an assimilatory segment addition process whereby a following single-vowel syllable [a] copies, in its onset position, the last segment from the preceding syllable. Examples of this process are given below in (2.15). The addition is shown in the IPA illustration. These examples show that the last segment of the previous syllable is invariably a high vowel (or its glide variant¹⁹) or one of the two nasal finals. This last segment is copied in the following syllable of the interjective particle [a] in the neutral tone:²⁰

¹⁸ See S. Wang (1994) for a discussion of phonotactic gaps in Chinese.

¹⁹ By the Glide-Formation Rule given later in (2.17), a high vowel becomes a glide when adjacent to a non-high vowel.

²⁰ See the section on Mandarin neutral tone in this chapter.

(2.14)

	IPA	Pinyin	
a.	zyg{ {c	huì a	会啊 '(pronoun) can!'
b.	vãJwa{æ	qù a	去呀 'Go!'
c.	zcy yc	hǎo a	好哇 'Good!'
d.	äkP pç	xíng a	行啊 'Will do!'
e.	mJcp pç	kàn a	看那 'Look!'
f.	t~ǰ ða	chī a	吃啊 'Eat!'
g.	tsl̥sa	zì a	字啊 Words!

Similar cross-syllable processes are quite rare. For instance, the alveolar nasal [n] is not copied in the second syllable even though that syllable starts with a low vowel [a] (e.g., *mián'ǎo* 棉袄 'cotton-padded coat', but not *mián nǎo*),²¹ indicating the unity and independence of the syllables in Mandarin. Interestingly, two reviewers of this book, both native American English speakers who are competent users of Mandarin-as-a-second-language, expressed difficulties with such lack of liaison between syllables in their Mandarin. One avoided using such terms as *mǐnÈnánÈyǔ* 闽南语 'Min dialect' since he would produce *mǐnÈnánÈnǚ* 闽南女 'Min girl' instead. The other has been heard to say *huāÈníng* when he meant to say *huānÈyíng* 欢迎 'welcome.' In both cases, the speakers applied English liaison rule to their production of Mandarin.²²

2.3.1 Consonants

In sharp contrast to the single phonological process across syllable boundaries, there are many such processes within the syllable. Mandarin phonemic vowels and consonants may have allophones: variants brought about by some contextual difference. In the Table of Consonants (2.4), we listed twenty consonants. However, all are phonemes—sounds that have minimal pairs to show that they are in contrast. Before we proceed to the Mandarin segmental phonological processes, an explanation of the concept of minimal pairs is in order here. Let us see an example from English. We know that the consonants [p] and [b] in English are phonemes because there are minimal pairs such as [pGt] 'pet' and [bGt] 'bet' to show that they are in contrastive distribution. The only significant difference between the two is that [p] is voiceless while [b] is voiced. The feature [voicing], therefore, is a contrastive feature in English. On the other hand, the aspirated

²¹ By Pinyin spelling conventions, if the non-initial syllable of a word begins with a vowel, that syllable is separated from the preceding one by an apostrophe (e.g., *mián'ǎo* 棉袄 'cotton-padded coat', and *xī'ān* 西安 'Xi'an, a city in Northwest China').

²² Note that in both cases, the liaison occurred between a previous consonant and a following vowel. In phonetic terms, *yǔ* and *yíng* begin with a vowel; the use of <y> in both is based on Pinyin spelling convention.

[pʃ] and the unaspirated [p] are variations of the same [p] and not two individual phonemes in English, even though both forms exist: [pʃ] in [pʃkt] 'pit' and [p] in [spkt] 'spit'. The crucial difference is that here [pʃ] and [p] are not in contrastive distribution, but in complementary distribution: the unaspirated stop occurs only after the sibilant [s] while the other occurs in syllable-initial position. In this case, the feature [aspiration] which differentiates the two is therefore not a contrastive feature.

Complementary distribution is one of the most important pieces of information used to determine whether or not two or more sounds go together as allophones of the same phoneme. Sometimes, however, complementary distribution alone is not sufficient. Consider the set of palatal consonants listed in (2.6), the Standard Table of Beijing Sounds and Syllables. These three sounds [t̪, t̪ʃ, ʃ] are heard only before high, front vowels (i.e., [i] and [ü]). On the other hand, there are three series of obstruents that are never heard before these two vowels, the dentals [ts, tsʃ, s], the alveopalatals [t̪ʃ, t̪ʃʃ, ʃʃ], and the velars [k, kʃ, x]; therefore, the palatals are in complementary distribution with all three. The question is to which one of these obstruent series should the palatals be assigned as allophonic variants?

One might be inclined to let the historical development of the palatals decide the matter. Most cases of these palatals have evolved from velars at an earlier time. The [x]-[t̪] and [k]-[t̪] correspondences²³ are clear comparing Cantonese [hakaw] 'shrimp dumpling' with Mandarin [äya-täyaw] 'shrimp dumpling'. One could, therefore, conclude that the palatals should be considered variants of the velars (i.e., [k]→[t̪], [kʃ]→[t̪ʃ], [x]→[t̪]). However, some of today's Mandarin palatals have developed from the dental series; and there is no way for modern Mandarin speakers to distinguish a palatal of the velar origin from one with a dental origin—at least not without philological study which is irrelevant for a synchronic study.

A better criterion for a decision on the allophonic membership of palatals rests on the concept of phonetic similarity. Presumably contemporary speakers of Mandarin feel palatals to sound more like alveopalatals than like either dentals or velars. One bit of evidence for this assumption is the fact that native Mandarin speakers are often heard replacing the English alveopalatals with the Mandarin palatals before high front vowels: [d\ip] 'jeep' becomes [t̪äip]; [tUiz] 'cheese' becomes [t̪äʃiz]; and [Uip] 'sheep' becomes [äip]. The evidence suggests that the native Mandarin speakers cannot produce the alveopalatals before a high front vowel or glide. On the other hand, the same is not true for velars and dentals. Both seem to induce fewer problems for the Mandarin speakers. The velar-initial English words [ki] 'key', and [gis] 'geese', and the alveolar-initial English word [sik] 'seek', for instance, are usually mastered without much problem. Thus, evidence from the ESL learning by native Mandarin speakers suggests that synchronically, the alveopalatals are the underlying phonemes from which the palatals are derived. If this is indeed the case, the phonological process of palatalization can be written as follows:

²³ The Mandarin consonant [x] in contemporary Cantonese is a glottal fricative [h].

(2.15) The Synchronic Palatalization Rule

alveopalatals → [+high, -back] / _____ [+vocalic, +high, -back]

Having given this tentative synchronic rule, we should mention that it is perhaps not unreasonable to assume that the phonemic origin of the palatals simply does not exist synchronically. The complementary distribution of the palatals versus the dentals, the alveopalatals, and the velars does not have synchronic significance, but is just a result of the historical development which palatalized some velars and dentals.

The following table gives a summary of all the consonants and glides in Mandarin. Non-phonemic ones are enclosed in braces (cf. (2.4)).

(2.16)

<i>Manner</i> \ <i>place</i>		<i>Labial</i>	<i>dental</i>	<i>alveolar</i>	<i>alveo-palatal</i>	<i>palatal</i>	<i>velar</i>
<i>Stop</i>	[+asp]	pʃ		tʃ			kʃ
	[-asp]	p		t			k
<i>affricate</i>	[+asp]		tsʃ		tʃ̥	(tǎʃ)	
	[-asp]		ts		tʃ̥	(tǎ)	
<i>fricative</i>	[+asp]	f	s		ʃ̥	(ǎ)	x
	[-asp]				Ê		
<i>approximant</i>	nasal	m		n			ŋ
	liquid			l			
	retroflex			ʈ			
<i>glide</i>		{w}				{y}{ÿ}	

2.3.2 Vowels

While it is evident that in Mandarin, there are only three outstanding allophones (the palatals) in consonants, there are more contextual variants in vowels. In fact, all six of the phonemic vowels have variations. A high vowel, for example, when occurring before or after another vowel (which is, incidentally, always [-high] in Mandarin), assumes the form of a glide; namely, [i] becomes [y], [u] becomes [w], and [ü] becomes [ÿ]. The processes can be captured by the following two mirror image rules. The first says that a high vowel becomes a corresponding glide if appearing before another vowel, and the second that a high vowel becomes a corresponding glide if appearing after a vowel.

(2.17) Glide Formation Rules

- a. [+vocalic, +high] → [-vocalic] / _____ [+vocalic]
 b. [+vocalic, +high] → [-vocalic] / [+vocalic] _____

Non-high vowels [H] and [a] also have variations. However, the variation does not involve categorical change; their allophones remain vowels regardless of environment. The mid vowel [H] has three context-induced allophones: [e, G, ɛ] whose distribution is given below:

- (2.18) a. /H/ → [e] / _____ i
 b. /H/ → [G] / [+high,-back]_____Σ
 c. /H/ → [ɛ] / _____ [+nas]

The rules say /H/ appears as [e] before the high-front-unrounded vowel [i] (see (a)), as [G] after [i] or [ü] and before the syllable boundary, denoted by 'Σ' (see (b)), as [ɛ] before the nasals (see (c)), and as [H] elsewhere. The following are examples that show the operation of these rules:

(2.19)

	IPA	Pinyin	
[e]	pey	bei	杯 'cup'
[G]	yG	ye	叶 'leaf'
	pyG	bie	鳖 'turtle'
	ÿG	yue	约 'approximate'
[ɛ]	ɛn	en	恩 'gratitude'
	tɛP	deng	灯 'light'
[H]	lH	le	乐 'happy'
	kH	ge	歌 'song'
	pJwH ²⁴	po	坡 'slope'

The low vowel [a], on the other hand, has just two context-induced allophones: [G] and [C]. Their distribution is illustrated below:

- (2.20) a. /a/ → [G] / [+high, -back] _____ n
 b. /a/ → [G] / _____ i
 c. /a/ → [C] / _____ P

²⁴ Many native speakers of Mandarin insert a high vowel [u] between the labial consonants [p, pJ, m, f] and [H]. After glide formation, the syllables become [pw H], [pJwH], [mwH], and [fwH] respectively (Pinyin <bo>, <po>, <mo>, and <fo>).

These rules spelt out in plain English are that the low vowel /a/ is raised to [ɤ] between the high front vowels [i] and [ü] and the alveolar nasal [n], or before the high vowel [i], but is backened before the velar nasal [ŋ]. Elsewhere, it remains as an [a]. Examples of these allophones are given below:

(2.21)

	<i>IPA</i>	<i>Pinyin</i>		
[ɤ]	yɤn	yan	烟	'smoke'
	tǎJyɤn	quan	圈	'circle'
[ɤ]	wɤy	wai	歪	'not straight'
[ɕ]	fɕp	fang	方	'square'
[a]	fan	fan	翻	'turn over'
	ma	ma	妈	'mother'

To summarize, all the Mandarin vowels, phonemic or phonetic (in braces), are given in the table below:

(2.22)

	<i>front</i>		<i>central</i>	<i>back</i>	
	<i>unround</i>	<i>round</i>		<i>unround</i>	<i>round</i>
<i>high</i>	i	ü			u
<i>mid</i>	(e) (ɤ)		(ɛ)	ɰ	o
<i>low</i>			a	(ɔ)	

2.3.3 Pinyin and IPA

Pinyin, literally 'spell sounds', is a Latin-letter-based alphabetic system that was initially developed in the 1950s in the People's Republic of China to replace Chinese characters. No sooner had the system been officially promulgated in 1958 than that purpose was abandoned. Thereafter Pinyin began to be used as a phonetic aid for the promotion of *Pǔtōnghuà* 普通话, literally, Common Language, the official name for Standard Mandarin.²⁵ Because it was developed as a writing system, Pinyin is not exactly a purely phonemic system. Rather, some symbols are phonemically oriented whereas others embody phonetic details.

Most consonant symbols are faithful representation of the phonemes of the language. There are three notable exceptions. As discussed in the last section, the palatal consonants [tǎ,

²⁵ See Chapter 1 for a description of the various names for Mandarin.

täɹ, ä] are not underlying phonemes, but allophones that appear only before high front vowels. They should, therefore, be missing from an inventory of phonemes in Mandarin. In Pinyin, they are treated as separate individual sounds from those from which they are derived and are represented by the symbols <j, q, x> respectively. Such treatment of these allophones is not without good reasons. As discussed previously, the phonemic origin of these allophones is still largely a mystery. This raises the question of which set of symbols to use for the representation of the palatals, those for the dentals, the alveopalatals or the velars? For instance, how should <ji> 'chicken' be written? Should it be written as <zi>, <zhi>, or <gi>? Besides, as mentioned earlier, it is perhaps not unreasonable to assume that the phonemic origin of the palatals simply does not exist synchronically. If this assumption is valid, it makes sense to represent these palatal sounds separately even from a phonemic point of view. In terms of dialect division, the palatalization serves to identify dialects of the north which have this feature from dialects of the south which generally lack it. Without representing the palatals separately, this important distinction between Mandarin in the North and Non-Mandarin Chinese dialects in the South would not be represented.

Compared with the consonant symbols, the vowel symbols of Pinyin are more faithful representations of phonemes in Mandarin. In (2.5), we saw that Mandarin has six phonemic vowels [a, o, ɤ, i, u, ü], and in Section 2.3.2, we learned, in addition, that two non-high vowels have allophones, and the high vowels have glide counterparts. Pinyin has six vowel symbols <a, o, e, i, u ü> and two glide symbols. They cover all the phonemic vowels, but not all the allophones. In particular, we have Pinyin <a> to stand for the allophones [ɤ] between [i] and [n], [ɤ] before [i], [ɕ] before [p], and [a] elsewhere. For the phonemic vowel /ɤ/, Pinyin has [e] before the high vowel [i], [ɤ] after the high vowel [i] and before the syllable boundary, [ɛ] before the alveolar and velar nasals, and [ɤ] elsewhere. For the three glides [w, y, ʏ] which are de-vocalicized high vowels, Pinyin has just two symbols: <y> and <w>. The former, <y>, is also the symbol for the de-vocalicized high front umlaut vowel [ʏ] when used in conjunction with a following <u> symbol (e.g., <yuan> [ʏɤn] 'circle').

In addition to not completely matching phonemic or phonetic information at the individual sound level, Pinyin differs somewhat from a strictly linguistic representation of the syllable. From (2.6) in the section on syllables, we learned that a syllable never has two high vowels adjacent to each other. Whenever two vowels appear together, one must be a non-high vowel, or whenever two high vowels are present, they would flank a [-high] vowel. In Pinyin representation, when the initial position is not empty (i.e., not a zero initial), the syllable final [yow] is <iu> (e.g., *līu* 溜 'to flee'), and the syllable final [wey] is <ui> (e.g., *tuī* 推 'to push'). In both cases, the middle non-high vowel is omitted. When there is a zero initial, however, this vowel is shown (e.g., *yǒu* 友 'friend' and *wěi* 尾 'tail'). Also shown is the glide formation of the initial high vowels. At the syllable level, another piece of information that is not codified in Pinyin concerns the umlaut vowel [ü]. Where the information is predictable from context, the diacritic is not shown in Pinyin. After the palatals <j, q, x>, for instance, the <ü> is simply <u> (e.g., *jū* 橘 'orange', *qū* 区 'region', and *xū* 需 'need'). When [ü] begins the syllable, it is represented with a <y> followed by <u> (e.g., *yū* 愚 'stupid', *yuán* 圆 'circle', *yuè* 月 'moon'). In fact, <ü> is spelt only after the initials <l>

and <n> (e.g., *nǚ*, 女 'woman,' *lǜ* 绿 'green', and *lūe* 略 'omit').

At the syllable level, Pinyin representation does not correspond to the phonemic information in yet another area: the representation of the syllables formed by only one syllabic consonant. As mentioned in Section 2.2.2.1, Mandarin has seven syllabic consonants: the sibilants alveolars and alveopalatals [tsʝ, ts, s, tʃʝ, tʃ, ʃ, ɛ̃]. These consonants provide substance for both the initial consonant and the peak vowel positions in the syllable structure. Pinyin uses the symbol <i> to represent the voiced part of the syllable, so that these seven syllables are represented with both a consonant and a vowel symbol: <zhi, chi, shi, ri, zi, ci, si>. Such representation does not create problems for native speakers of Mandarin learning Pinyin, but may cause trouble for Chinese-as-a-second-language learners who may mistake the vocalic part of the syllable for the high front unround vowel [i] found in *tī* 踢 'to kick' and *mǐ* 米 'rice'.

Some people may argue that the vocalic part in the syllables of syllabic consonants is actually a vowel that is in complementary distribution with the high front vowel [i], and therefore, the use of the <i> in both cases is justified. This usage signifies that the vowel of the syllabic consonant is just an allophone to the underlying vowel [i]. Such a relation, however, cannot be established. We learned previously in Section 2.3.1, that the dentals, alveopalatals, and velars are in complementary distribution with the palatals. The complementary distribution is quite clear cut, and there is no reason to believe that the complementary distribution has just one exception: the syllables with a single high front vowel [i]. Additionally, some dentals before high front vowels became palatals historically. Treating the vowel in the syllabic-consonant syllables as underlyingly a high front vowel would leave unanswered the question why these consonants have not become palatals. A third argument against assuming the vocalic part of these syllables to be the high front vowel is that the vowel part in the dental syllables and those in the alveopalatal syllables do not have the same value²⁶; it is not clear if they are the same vowel.

Since the oldest system of romanization for Chinese was introduced by Sir Thomas Wade in 1859 and developed by Herbert Giles in his 1912 dictionary, some half a dozen alphabetic systems have been developed and used for writing and teaching Chinese. In addition to varied choices of symbols for certain sounds, one major difference among them lies in how they treat phonemic and phonetic information. In particular, an allophone represented separately from its phoneme in one system may not be so in another. For instance, in the Wade-Giles system, the palatals [tä, täʝ, ä] are treated as if they are allophones of the alveopalatals²⁷ since both sets share the same set of symbols as the retroflex [tʃ, tʃʝ, ʃ]; namely <ch, ch', sh> respectively. In Pinyin, on the other hand, symbols for the former are <j, q, x,> while those for the latter <zh, ch, sh>. Another example is that the Wade-Giles system does not distinguish between the aspirated affricates [tʃ, tʃʝ] in the primary symbols; both are <ch>, although they are distinguished by a secondary symbol, an apostrophe: <ch, ch'>. In Pinyin, they are represented distinctively in primary symbols as <zh> and <ch>. In sum, segmental phonology sometimes is and sometimes is not encoded in the alphabetic systems of Chinese.

²⁶ They do not rhyme with each other. Neither do they rhyme with [i].

²⁷ Incidentally, this is the position mentioned in (16).

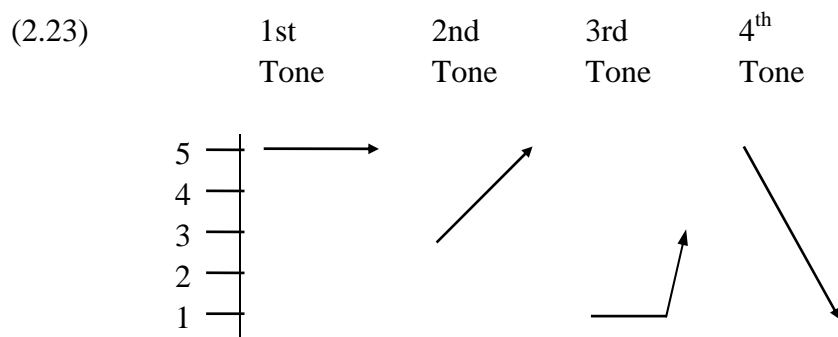
2.4 The Tones

In the last section, we looked at segmental processes in Mandarin. Now let us turn to the supersegmentals, in particular, the tones of Mandarin. That Mandarin is a tone language is one of its most well-known features. But what is a tone? How does it work in language? In particular, what are tones like in Mandarin? These are among the questions that will be explored in this section.

The understanding of tone relies on knowledge of another linguistic entity called pitch. Acoustically, pitch can be measured in terms of fundamental frequency (F_0). Phonetically, the production of pitch involves the tensing of the laryngeal muscles. All languages manipulate pitch to encode meaning. However, while all languages do so at the phrase or sentential level which is called intonation, only some languages use pitch variation at the word/morpheme level. Languages that vary pitch at the word/morpheme level are said to be tone languages. A very simple example of the linguistic use of pitch at the word level can be found in Nupe, a language spoken in Nigeria. In this language, the syllable *ba* means 'to count' if said with a low pitch, but it means 'to be sour' if said with a high pitch.

2.4.1 Basic Tones

Not all tone languages use pitch in the same simple way as the Nupe example suggests. Pitch at the word or morpheme level is used in a more complex way in Mandarin, as it is in many Asian languages. In Mandarin, it is not just a matter of the height of the pitch varying to yield different level tones, but rather the pitch may go upward or downward to produce different contour tones. Several linguistic analyses have been developed in the past to describe tones. Some are more theoretically accomplished, more precise and insightful, while others are less so. One of the most used but less theoretical descriptions of tones was the Scale of Five Pitch Levels, developed by the late Yuen Ren Chao, the most well-known linguist of Chinese. On this Scale, the four basic Mandarin tones can be represented as here:



What the Scale does is to divide the pitch range of a person engaging in normal speech into five levels, with Level 5 representing the highest pitch, and Level 1 the lowest. Traditionally, tones in Mandarin are each given a number for ease of reference. The First Tone, the only level tone of the language, starts at the highest point on the Scale and continues at that same pitch for a designated period of time. The Second Tone, a rising tone, starts somewhere in the middle of the pitch range and rises upward to the highest point. In the production of the Third Tone, the speaker attempts to

start at the lowest pitch, continues at that low pitch for a little while, and then raises it somewhat before finishing. In producing this tone, the speaker may unintentionally start a little higher than the intended low pitch target due to the physiological constraints which make it hard for the speaker to reach the low target right away. Traditional descriptions which assume the third tone's value to begin with a falling pitch have failed to consider this. The Fourth Tone is falling and covers all the pitch range. It starts at the highest pitch level and drops to the lowest. Based on the measurement on Chao's Scale, the First Tone can be said to have the value of (55), indicating that the tone starts at the pitch level of 5 on the Scale and finishes at the same level. For the same reason, the rest of the four tones can be said to have the value of (35), (113) and (51). The four Mandarin tones as given here are the citation tones. That is, they are tones produced in isolation or in a final position where no other tones follow. In Pinyin, these four tones are represented by diacritics over the vowel, or if there are more than one vowel, over the lowest one. These four diacritics are shown below over the syllable <ba>.

(2.24)

	<i>Pinyin</i>		
<i>1st Tone</i>	bā	八	'eight'
<i>2nd Tone</i>	bá	拔	'pull out'
<i>3rd Tone</i>	bǎ	靶	'target'
<i>4th Tone</i>	bà	爸	'father'

In connected speech, these basic Mandarin tones often undergo context-induced modifications called tone sandhi. The best-known tone sandhi processes in the language involve the Third Tone (113). This tone, which is also phonetically the longest of the four in citation form (L Feng 1985), becomes a high rising tone similar to the second tone²⁸ (35) when appearing before another Third Tone²⁹. However, it becomes a low level tone (11), a 'half third' in Chao's terminology, when appearing before other basic tones. In other words, the Third Tone never takes on the citation form if there is another tone following it. The Fourth Tone also undergoes sandhi. This tone, when followed by another tone, is modified into (53) from the citation form of (51)³⁰. Similarly, the Second Tone has been found to alter its shape in connected speech. It becomes high

²⁸ It is not conclusive at the present time whether the modified Third Tone is completely identical to the Second Tone in value. According to Shen (1990), for instance, the second Third Tone is higher when following an inherently rising tone (i.e., a Second Tone) than a rising tone generated by the Third Tone Sandhi (p. 285), which indicates that the two rising tones are not quite the same.

²⁹ Tone sandhi as presented in this section is always true on an isolated two-syllable sequence. However, when the string gets longer and involves more than two syllables, the picture becomes more complex. Such complexity has given rise to sentence phonology, which will be briefly discussed later in this chapter.

³⁰ Traditionally, the Fourth Tone was found to undergo this change only before another Fourth, but this understanding was based on imprecise auditory impression. Shih's (1986) Bell Laboratory report indicated that the Fourth Tone never reaches the low target when followed.

level when sandwiched by the First or the Second in front and another tone after it. The only tone that, for a long time, was not found to undergo modification is the First Tone. Is there really a linguistic gap here? Recent studies by H. Lin (1992, 1996, and 1998) have suggested otherwise.

In these studies, we found that the perceived absence of modification on the First Tone is a result of the limitation of the theoretical tools used in the description of tones. Traditional phonological analyses of tones have looked exclusively at the direction along which the pitch moves: upward, downward or just steadily ahead. Theoretically, they ignored treating the timing along which the movement occurs. We suggest an analysis that takes the timing into account. In this analysis, each Mandarin tone is represented by three tonemes, each toneme representing, relatively speaking, one third of the time used to utter the tone in citation. Using the tools developed in the non-linear phonological theory in which tones are represented by a sequence of level tonemes, the four Mandarin tones are represented in citation as HHH, MHH, LLM and HML. The symbols 'H, M, L' represent high level, mid level and low level tonemes respectively. This representation of the four Mandarin tones takes into account the fact that, phonologically speaking, all Mandarin tones (or syllables) have the same relative length³¹. One important fact about tones in Mandarin is that the citation form is about twenty-two percent longer than the same tone in non-final position (L. Feng 1985)³². This figure, translated into relative toneme representation, can be construed as the shortening of a non-final tone by the deletion of the last toneme. In this way, the four Mandarin tones should have the shapes of HH, MH, LL, HM, respectively, when not appearing alone or in final position. As we will show below, this treatment can explain the previously mentioned Half-Third Sandhi, Fourth Tone Sandhi and Second Tone Sandhi by making the following stipulation:

(2.25)

In connected speech, within the metrical foot of two syllables, a toneme at the edge of a basic tone is deleted if (a) adjacent to another tone, and (b) not in a final tone.

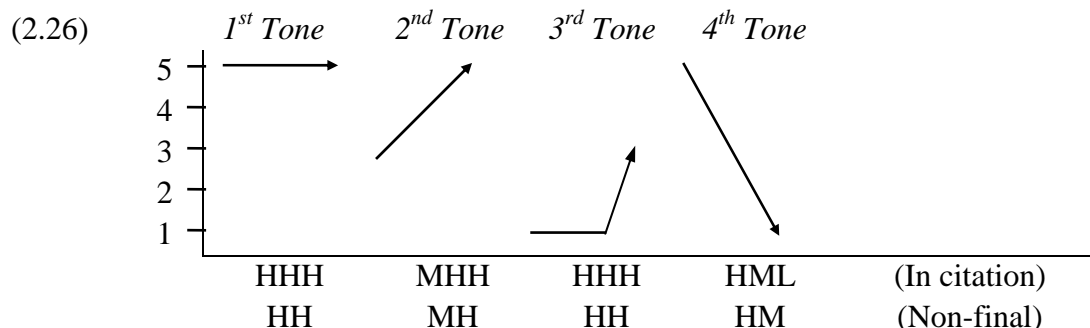
One advantage of treating the timing dimension in Mandarin tonal analysis is that it yields an explanation as to why the First Tone does not undergo sandhi change. We find that it is not that the First Tone does not change, but that the change is not detectable in traditional theory of tonal representation. The First Tone, just like the other tones, is modified in its length when another tone follows: Just as the Fourth Tone loses the last toneme to change from HML to HM (i.e., from 51 to 53), the First Tone changes from HHH to HH. But because the traditional tonal theories do not encode length, this tonal change in the First Tone is not treatable and the fact that the First Tone undergoes the same change as the Fourth Tone (and the Second and Third) is concealed.

³¹ Acoustically, the Third Tone is the longest, then followed by the Second, the First and the Fourth. However, the difference is insignificant at the rule-governed phonological level.

³² Feng's findings are that, in non-final position, the average tone lasts 235 milliseconds while in final position, it lasts 301 milliseconds (p. 177).

Another advantage is that now we can pinpoint the reason why a Third Tone has to change when another Third Tone follows. The reason that emerges from our theory is dissimilation, a very common process found in the world's languages. When a Third Tone is followed by another Third, by (2.25), it loses its last toneme. As a result, there appears a series of more than two low tonemes (i.e., LL LLM), a sequence that is perhaps physiologically hard to produce. Adjustment thus has to ensue, and the result is that the tail of the first tone is raised.

The following summarizes our theory of Mandarin basic tone with reference to Chao's Scale:



In other words, all the four basic tones have sandhi processes which can be captured, first of all, in the following rules (which are essentially the same rule—deletion of a final toneme). Note that the symbol 'X' stands for any toneme, H, M, or L.

(2.27)

	<i>In Citation</i>		<i>In Normal Speech</i>
a. 1^{st} Tone Sandhi:	T T / \ / \	→	T T /\ / \
b. 2^{nd} Tone Sandhi:	HHH XXX T T / \ / \	→	HH XXX T T /\ / \
c. 3^{rd} Tone Sandhi:	MHH XXX T T / \ / \	→	MH XXX T T /\ / \
d. 4^{th} Tone Sandhi:	LLM XXX T T / \ / \	→	LL XXX T T /\ / \
	HML XXX		HM XXX

In fast speech, the Second Tone goes further by losing the left toneme when it is, at the same time, preceded by another tone³³ (Y. R. Chao 1968).

³³ Chao only detected the high level Second Tone when the preceding tone is a First or a Second tone. However, there is reason to believe that the same is true when the preceding tone is a Third or a Fourth. On the one hand, the

(2.28)		<i>In Normal Speech</i>	<i>In Fast Speech</i>	
b'. 2 nd Tone Sandhi:	T T T			T T T
(continued)	/ \ / \ / \	→		/ \ / \
	XXX MH XXX			XXX H XXX

The Third Tone, if followed by another Third, now gives rise to a problem—the output contains a string of more than two low tonemes. A repair mechanism is therefore activated whereby an edge toneme is raised (See Footnote 28) to break the monotony:

(2.29) c'. 3 rd Tone Sandhi:	T T		T T		T T
	/ \ / \	→	/ \ / \	→	/ \ / \
	LLM LLM		LL LLM		LM LLM

2.4.2 Neutral Tone

That Mandarin does not favor a sequence of low tones is also found to be significant in the case of the neutral tone. But before we get to the low tone issue, let us become acquainted with the Mandarin neutral tone. This is unfortunately not an easy and straightforward task due to the fact that there is still a great deal of controversy concerning what exactly the neutral tone is. In this section, we will attempt to present briefly the crucial issues and work out a coherent description of the neutral tone.

That the neutral tone is short and weak is well recognized. It is somehow derived from the four basic tones; it is called the neutral tone because, in unstressed positions, any of the four basic tones can lose their inherent tone and be 'neutralized' into this short and weak tone. This tone neutralization process has been active for quite a long time: Many lexical items in the language—mostly function words which almost never appear in stressed positions—have become fossilized Neutral tones, their tonal origin hard to trace from a synchronic perspective. Among words with neutral tones whose tonal origin is traceable, some neutral tones are compulsory while others are optional due to a number of factors unclear at the present time (See Lin 2001 for an explanation).

Phonetically, the neutral tone has been found to have not just one value but a number of values depending on which basic tone it follows. By auditory impression, its phonetic values have been found to be the following (S. Q. Qi 1956, T. Lin 1962, Y. R. Chao 1968 and Z. J. Wu 1985):

(2.30) *Basic Tone+Neutral Tone*

(55)(3)
(35)(3)

fact that the medial tone is the shortest is well-documented (e.g., Chao 1968, Yan and Lin 1988, and Shen 1990b), and on the other, the characteristic of a level tone may have been concealed by transition from the lower offset of the preceding Third or Fourth tone.

(11)(4)

(51)(1)

After the First, Second and Third Tone, it is relatively high, that is, (3) or (4) on Chao's scale, and after the Fourth it is low or (1) on it. Furthermore, the neutral tone has been found to be contoured in experimental studies using phonetic/acoustic instruments (Dreher and Lee 1966):

(2.31) *Basic Tone+Neutral Tone*

(55)(41)

(35)(31)

(11)(23)

(51)(21)

It falls after the First, Second and Fourth tones but rises after the Third. If these are its phonetic values, what is the phonological—the phonemic or underlying—value of the neutral tone? One of the best known positions is that the neutral tone does not have an underlying value, but derives its surface tone from the preceding tone. This position fails to capture the native speaker's intuition that the neutral tone is generally low after the First, Second and the Fourth, but high after the Third. The well-known linguist Chin Chuan Cheng (1973) once remarked:

'The refined acoustic details [about the neutral tone] perhaps do not necessarily represent the native speaker's knowledge. The speaker's aim perhaps is to produce the neutral tone low after the first, second, and fourth tones and higher after the third tone.' (p. 56)

A closer look at the measurements acquired in Dreher and Lee's study suggests the same dichotomy, in that the neutral tone goes low (and falls) after the First, Second and Fourth Tones, but high (and rises) after the Third. All these considerations point to a classic case of what is called complementary distribution, in which one of the variants is the underlying form—the basic form. It would not take long for any linguist to determine that the low tone after the First, Second and Fourth is the underlying one since the environment in which it occurs is more complicated and less predictable.

Interestingly, our theory that the neutral tone is fundamentally a low tone solves yet another problem. If we take the position of the traditional belief that the neutral tone is toneless on its own but derives its tonal shape from the basic tone before it, a problem arises when more than one neutral tone appears after the basic tone. Although phonetically we know the values of the neutral tone after the four basic tones, the question is what value should we assign to the neutral tone that does not follow a full tone but another neutral tone? Such neutral tones are found in Mandarin phrases such as these:³⁴

³⁴ Refer to the chapters on morphology and syntax for an explanation of the grammatical markers PERF and DE(a).

- (2.32) a. tuī shang qu le de rén 推上去了的人
 push up go PERF DE(a) person 'person pushed up'
- b. tái shang qu le de rén 抬上去了的人
 carry up go PERF DE(a) person 'person carried up'
- c. zǒu shang qu le de rén 走上去了的人
 walk up go PERF DE(a) person 'person who has ascended'
- d. sòng shang qu le de rén 送上去了的人
 send up go PERF DE(a) person 'person sent up'

All four phrases begin with a monosyllabic word, one that carries the First (a), the Second (b), the Third (c) and the Fourth Tone (d) respectively. All of the four syllables between the first word and the last (*rén* 人 'person') are in the neutral tone. Phonetically, except for the first neutral tone, all the other three are strictly low tones. The question here is where these three neutral tones get their low tonal values now that they are not adjacent to a basic tone? What we have argued so far boils down to the following conclusion: that the Mandarin Neutral tone is fundamentally a low tone. In the notation of current theoretical phonology, the low neutral tone can be represented as follows:

(2.33) The Underlying Representation of the Mandarin Neutral Tone

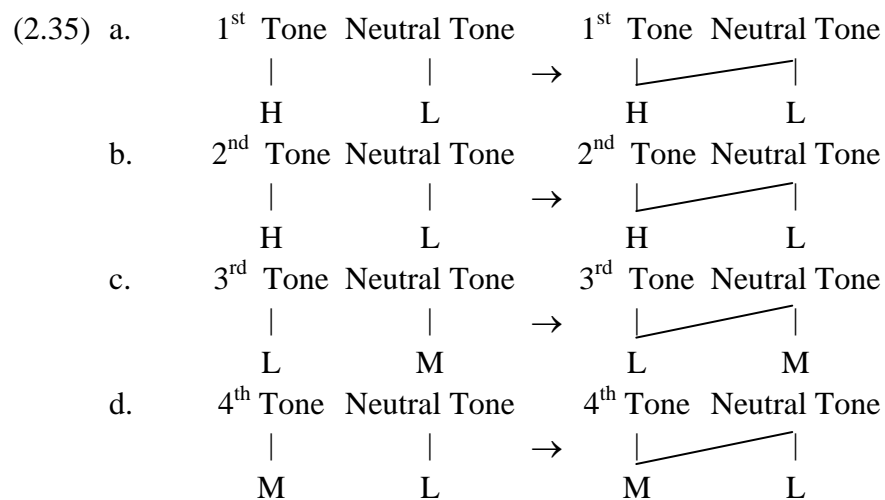
T
|
L

Now that we have pinpointed the basic value of the Neutral tone, we can look at the neutral tone after the Third Tone to see what causes the low neutral tone to become high after the Third Tone. Let us juxtapose the two tones:

- (2.34) 3rd Tone Neutral Tone
- | |
LL L

Here we find the same low tone sequence as found in the case of the Third Tone Sandhi. We learned earlier that Mandarin avoids sequences of low tonemes, and that the low tone sequence created by juxtaposing two Third Tones triggers some repair process that raises the tail of the preceding Third Tone. If the language does not like sequences of low tonemes when the Third Tone is concerned, there is no reason to expect it to allow the low tone sequence shown above in (2.34). The only difference between the Third Tone Sandhi in (2.29) and the case here lies in that the repair affects the first syllable in that case, but the second in this one. The difference in the repair, however, should not be difficult to understand, since in both cases, it is the prosodically weaker (and therefore less stable) tone that undergoes the change.

In brief, we believe that the Mandarin neutral tone, which is induced by the lack of stress, is essentially a short low tone. It dissimilates after the Third Tone by rising in pitch. Interestingly, we can now account for the varied surface values of the neutral tone in a consistent manner. They are all the result of tone spreading, which can be elucidated by a line linking the preceding toneme and the neutral tone, as shown below:



After spreading, the resultant neutral tone after the First, Second and the Four Tones becomes HL, HL, and ML—falling tones, respectively, but an LM, a rising tone, after the Third. These correspond to the results of the instrumental study by Dreher and Lee (1966) mentioned earlier.