17 Tones, Tonal Phonology, and Tone Sandhi

JIE ZHANG

1 Introduction

Chinese dialects are well known to be tonal, in the sense that the pitch with which a syllable is uttered can cue meaning differences. Examples that illustrate the contrastive status of tones abound in virtually all Chinese dialects, and the ma55/ma35/ma213/ma51 “mother/hemp/horse/to scold” quadruplet for Standard Chinese (SC), made popular by introductory linguistics textbooks, provides only a glimpse of this robust phenomenon.

The differences between Chinese tone languages and African tone languages were noted as early as by Pike (1948). Wan and Jaeger (1998: 426–427) aptly summarized the differences as follows: first, tones in Chinese are typically associated with individual syllables and predominantly serve a lexical function, while tone patterns in African languages are often associated with polysyllabic phonological words and serve grammatical functions; second, Chinese languages generally have a contour-based system with multiple contour tones in the tonal inventory, while African languages usually have a register-based tone system with two or three tonal levels.

Added to these properties is the fact that Chinese languages often have complex patterns of tone alternation caused by adjacent tones or the prosodic/morphosyntactic environment in which a tone appears, commonly referred to as tone sandhi. Two examples of tone sandhi from SC and Taiwanese are given in (1): in SC, the third tone 213 becomes 35 before another 213 (1a), and in Taiwanese, a tone undergoes regular changes whenever it appears in non-XP-final positions (1b) (Chen 1987; Lin 1994; Simpson, Chapter 18, this volume).

© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.
(1) Tone sandhi examples:
   a. Tonally induced tone sandhi – SC “third-tone sandhi”:
      \[213 \rightarrow 35 / \_\_ 213\]
   b. Positionally induced tone sandhi – Taiwanese “tone circle” (open or sonorant-closed syllables):
      \[51 \rightarrow 55 \rightarrow 33 \leftrightarrow 24 \text{ in non-XP-final positions}\]

These unique properties of Chinese tones determine that they have played unique roles in the development of phonological theory, from the autosegmental nature of phonological features (Wang 1967; Woo 1969) and the feature-geometric representations of tones (Bao 1990, 1999; Chan 1985, 1991; Chen, M. 1996; Chen, T.-Y. 2010; Duanmu 1990, 1994; Yip 1980/1990, 1989, 1995) to the relevance of phonetics to phonological markedness (Jiang-King 1999; Zhang 2002, 2004, 2007; Zhang and Lai 2010) and how variations and exceptions can be treated in phonology proper (Zhang and Lai 2008; Zhang et al. 2009, 2011). The primary goal of this chapter is to discuss in further detail the typological characteristics of Chinese tone patterns and outline the contributions that the studies of Chinese tones have made to phonological theory. Secondarily, the methodological issue of how to best study Chinese tone patterns in the future is also addressed.

2 Typological characteristics of Chinese tones

Chinese can be generally classified into seven dialect groups: Northern, Wu, Min, Yue, Kejia (Hakka), Xiang, and Gan (Yuan 1983). There are also distinct dialects within each dialect group. Aside from historical and geographic factors, there are many linguistic criteria for dialectal grouping, from phonological, to syntactic, to lexical (Norman 1988). With respect to tones, different dialect groups tend to have different types of tonal inventories depending on whether they have kept the historical obstruent codas, and often have distinct properties in their tone sandhi systems. Dialects spoken in the South, in particular Min, Yue, Kejia, and Gan, are phonologically more conservative and have typically maintained more tonal contrasts than the Northern dialects.

2.1 Characteristics of tonal inventories

Typological studies on tone in Chinese dialects and elsewhere have identified a number of implicational relations for tonal inventories, including: (i) contour entails level (Maddieson 1977; Zhang 2002), (ii) complex contour entails simple contour (Maddieson 1977; Zhang, 2002), (iii) rising entails falling (Zhang 2002); (iv) rising-falling entails simple rising (Cheng 1973); and (v) low level entails high level (Jiang-King 1999). However, a careful look at the tonal inventories in Chinese dialects shows that most of these generalizations have exceptions. For instance,
Jiang (1999) cites Pingyao (Hou 1980) for (i), Yangqu (Meng 1991) for (iv), Taiping (Zhang 1991) for (v), and Zhang (2002) cites Zengcheng (He 1986) for (iii). On the other hand, there do seem to be strong statistical tendencies when the tonal inventories of different dialects are considered as a whole. Jiang-King (1999), in a typological survey of 40 Chinese dialects, found that: (i) high level is more common than mid level, which is in turn more common than low level; (ii) simple contours with small pitch excursions are more common than those with large excursions; (iii) simple contours are more common than complex contours; and (iv) falling-rising contours are more common than rising-falling contours. With our increased understanding of tonal typology, it is likely that the implicational relations are better stated as statistical tendencies rather than sound laws.

Another issue of contention on tonal inventories is the number of contrastive level tones that can be found in a language (see Yip 1980/1990, 2002; Duanmu 1996; Jiang 1999). Chinese dialects with three or four level tones are relatively common. Reports of five level tones, though rare, can also be found; Jiang (1999) cites Rongxian (Zhou 1987) and Hengxian (Bi 1994) – two dialects spoken in Guangxi Province – as examples, and such systems exist in the Miao-Yao (Chang 1953) and Dong (Shi et al. 1987) languages spoken in Mainland China as well. The maximum number of contrastive tone levels has important connotations on tonal representations – an issue we return to in Section 3.

In dialects that have preserved the obstruent codas (Wu, Min, Yue, Kejia, and Gan), there is invariably a reduced tonal inventory on this type of syllables (CVO) as compared to open or sonorant-closed syllables (CVV, CVR): the number of tones that can appear on CVO syllables is usually one or two, occasionally three, while CVV and CVR syllables always carry a larger number of tonal contrasts. Illustrative examples from Cantonese (Yue), Shanghai (Wu), and Fuzhou (Min) are given in (2). This finds correspondence in historical Chinese as well: there were only two tonal categories on CVO syllables (\textit{yin ru} and \textit{yang ru}), while there could be a maximum of six on CVV and CVR (\textit{yin} and \textit{yang} cross-classified with \textit{ping}, \textit{shang}, and \textit{qu}). It is also often noted that the tones on CVO syllables are either level or have less pronounced pitch contours than tones in CVV/CVR syllables (e.g., Duanmu 1990, 1994; Yip 1995; Zhang 2004), and this fact has featured prominently in the debate on the nature of the Tone Bearing Unit (TBU), which is discussed in Section 3.

(2) Tonal inventories on CVV/CVR and CVO syllables:
\begin{tabular}{lll}
 & CVV/CVR & CVO \\
Cantonese (Matthews and Yip 1994) & 55, 33, 22, 35, 21, 23 & 5, 3, 2 \\
Shanghai (Zhu 2006) & 52, 34, 14 & 4, 24 \\
Fuzhou (Liang and Feng 1996) & 44, 53, 32, 212, 242 & 5, 23 \\
\end{tabular}

2.2 Characteristics of tone sandhi

The typology of tone sandhi in Chinese languages is well studied. This is due to collective efforts from both Chinese dialectologists who publish careful
descriptions of the dialects\textsuperscript{3} and theoretical linguists who make generalizations of the patterns observed in these descriptions. Influential dissertations by Yip (1980), Wright (1983), Bao (1990), and Duanmu (1990), typological works by Yue-Hashimoto (1987) and Bao (2004), as well as a long line of work by Chen (e.g., 1991, 1992, 1996) that culminated in his seminal tome on tone sandhi (Chen 2000) have all made important contributions to our understanding of Chinese tone sandhi.

Crudely speaking, tone sandhi patterns in Chinese fall under two types: “right-dominant” and “left-dominant” (Yue-Hashimoto 1987; Chen 2000; Zhang 2007). Right-dominant sandhi, found in most Min, Southern Wu, and Northern dialects, preserves the base tone on the final syllable in a sandhi domain – the grammatical domain over which sandhi generalizations are stated – and changes the tones on non-final syllables. Left-dominant sandhi, typified by Northern Wu dialects, preserves the tone on the initial syllable. Zhang (2007) argues that there is an asymmetry in how the sandhi behaves based on directionality, in that right-dominant sandhi tends to involve local or paradigmatic tone change, while left-dominant sandhi tends to involve the extension of the initial tone rightward. This asymmetry can be seen in the right-dominant examples in (1) and the left-dominant examples in (3): both the tonally induced “third-tone sandhi” in SC and the positionally induced “tone circle” in Taiwanese involve local paradigmatic tone change, while the sandhi in Changzhou (Northern Wu) spreads the tone on the first syllable across the entire sandhi domain (Wang 1988).

(3) Left-dominant sandhi – Changzhou:

\[
\begin{array}{ccc}
\sigma_1 & \sigma_1 \sigma & \sigma_1 \sigma \\
55 & 33-33 & 33-33-33 \\
13 & 11-33 & 11-33-55 \\
45 & 45-55 & 45-55-55 \\
523 & 55-23 & 52-22-33 \\
24 & 11-24 & 11-11-24 \\
\end{array}
\]

Different dialects vary greatly in the complexity of their tone sandhi patterns, from Cantonese and many Northern dialects that have relatively few sandhi changes to Min and Southern Wu dialects in which virtually any tonal combinations will trigger sandhi. The complexity also lies in the phonetic arbitrariness of the sandhis. Typological studies have found that in paradigmatic changes, despite the relative commonality of the phonetically expected contour tone simplification, tone level assimilation, and contour dissimilation (e.g., Rise-Rise \rightarrow Fall-Rise), the opposite patterns are also attested (Yue-Hashimoto 1987; Chang 1992; Chen 2000); in left-dominant tone extension, there are often nuanced phonetic changes that cannot be predicted by a straightforward tone mapping mechanism. For example, in the Fuzhou tone sandhi pattern on CVV/CVR syllables in (4) (Liang and Feng 1996), the non-final syllable undergoes sandhi, and the sandhi involves neutralization of the base tones 44, 212, and 242. But the tones that these base tones neutralize to seem largely arbitrary: although the changes in 44 may be considered as assimilation of the offset of 44 to the onset of the second tone (e.g., 44 \rightarrow
53 / __ 2), no such motivation can be found for the behavior of 212 or 242. The sandhi tones of 53 and 32 also seem largely arbitrary phonetically: tonal assimilation and contour reduction may account for the change from 53 to 44 before 44, but not other changes. The Changzhou example in (3) also shows that tone extension cannot account for certain details in the sandhi pattern such as the change in pitch height of an underlying tone 55.

(4) Fuzhou tone sandhi on CVV/CVR syllables:

<table>
<thead>
<tr>
<th></th>
<th>σ₂</th>
<th>σ₁</th>
<th>44</th>
<th>212</th>
<th>242</th>
<th>53</th>
<th>32</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>44-44</td>
<td>53-212</td>
<td>53-242</td>
<td>44-53</td>
<td>53-32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>212</td>
<td>44-44</td>
<td>53-212</td>
<td>53-242</td>
<td>44-53</td>
<td>53-32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>21-44</td>
<td>44-212</td>
<td>44-242</td>
<td>21-53</td>
<td>24-32</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A major cause for the phonetic arbitrariness of the synchronic tone sandhi systems is diachronic sound changes that wiped out the phonetic motivations of the sandhis due to tone’s fickle phonetic nature, as argued by Chen (2000) and Zhang (2010). This is evidenced by the fact that closely related dialects often have different phonetic manifestations of the same historical sandhi. For instance, the “third-tone sandhi” in SC in (1) corresponds to the historical sandhi pattern shang → yang ping / __ shang, which dates back to the sixteenth century (Mei 1977) and has manifestations in many other Northern dialects; but in Jinan, it is 55 → 42 / __ 55 (Qian and Zhu 1998), and in Taiyuan, it is 53 → 11 / __ 53 (Wen and Shen 1999). The Taiwanese “tone circle” pattern in (2) also finds correspondence in other Southern Min dialects, such as Longxi, whose sandhi pattern is shown in (5) (Chen 2000). The phonetic arbitrary nature of tone sandhi poses particular challenges for using tone sandhi to motivate the formal representations of tones as well as the theoretical analyses of the sandhi patterns themselves, as we will see in Sections 3 and 4.

(5) Longxi tone sandhi:

52 → 14 → 33 ← 313

Another level of tone sandhi complexity lies in the phonological opacity (Kiparsky 1973) often manifested in the sandhi patterns. In the Taiwanese example in (1b), four of the five tones in the tonal inventory on CVV/CVR are involved in a circular chain shift. In Fuzhou (4), chain shift also occurs: 32 → 44 → 53 → 21 / __ 212, 242; 44 → 53 → 32 → 24 / __ 32. These patterns pose analytical challenges for constraint-based Optimality Theory (Prince and Smolensky 1993/2004): circular chain shift has been shown to be incomputable by a “conservative” OT grammar that uses only IO-faithfulness and markedness constraints (Moreton 2004), and
regular chain shifts also require additional mechanisms such as constraint conjunction to be captured (Kirchner 1996). We return to this issue in Section 4.

3 Tonal representation and the TBU

The issue of the formal representation of tone is highly contentious, and phonologists’ arguments for any specific theory typically come from properties of tonal inventories and tone sandhi behaviors. The issue revolves around the following questions:

(6) a. What is the Tone Bearing Unit (TBU)? Is it the syllable, the rhyme, the sonorous portion of the rhyme, or the mora?
b. What are the primitive features of tone? How many level-tone features are needed, and are contour tones represented by unitary features or sequences of level features?
c. If there are different layers of tonal features such as Register (representing the overall pitch height) and Contour (representing TBU-internal pitch change), what is the geometric relation among the features – independence, dominance, or sisterhood?

Based on answers to these questions, a high falling tone over a syllable can be represented as one of the five representations in (7). The TBU in (7a)–(7d) is on the syllabic level (syllable, rhyme, or sonorous portion of the rhyme), while the TBU in (7e) is the mora. (7a) represents the contour tone as a single unit with a [+fall] feature (Wang 1967). (7b)–(7d) assume two levels of tonal representation: Register (High/Low) and Contour (high/low). These two features may be in an independence ((7b); Yip 1980/1990), dominance ((7c); Yip 1989, 1995), or sisterhood ((7d); Bao 1990, 1999) relation. (7e) represents the contour tone as a sequence of two level tones with no contour tone unit (Duanmu 1990, 1994). In (7b)–(7e), the Register and Contour features may take on different numbers of levels, but binary distinctions H/L and h/l are the most commonly assumed.4

(7) a. TBU
    [+[fall]] TBU
    h l
b. H c. TBU
    H h l
d. TBU e. TBU TBU
    r c H h H h
    H h

The issue of the maximum number of level tone contrasts mentioned earlier played an important role in the decision to represent tone with two binary fea-
tures, which can easily derive four contrastive level tones. Yip (1980/1990) expressed doubt about the existence of languages that truly have five contrastive level tones, yet Shi et al.’s (1987) phonetic study showed that Gaoba Dong does indeed have five level tones. Duanmu (1996) offered the possibilities of representing the fifth level tone as underlyingly toneless or relaxing the binarity requirement for the tone features, but did not pursue these arguments further with empirical support.

Whether the TBU is on the syllabic or moraic level is closely related to the decision to treat contour tones as either units or concatenations of level tones. Proponents of moraic TBUs, most notably Duanmu (1990, 1994), observed that the free distribution of contour tones in a Chinese dialect is often concomitant with longer rime duration, coda contrasts, and diphthongs in the dialect (e.g., SC, Fuzhou), while shorter rime duration and the lack of coda contrasts and diphthongs often give rise to greater contour restrictions and tone sandhies that wipe out contour tones (e.g., Shanghai, Suzhou). This correlation between rime properties and contour tone licensing can be easily captured by the bimoraicity of the longer rimes and the moraic licensing of each tone level in a contour tone. The predominance of level tones on the bimoraic CVO syllables, under this analysis, is due not to the obstruent codas’ inability to license tones phonologically, but to their inability to realize tones phonetically. Relatedly, Duanmu (1990, 1994) and Chen (2010) argued against the true existence of assimilation and dissimilation of tonal contours as a whole in tone sandhi processes on typological and phonetic grounds. However, Zhang (2004) showed that the correlation between rime properties, especially the sonorous rime duration, and contour tone licensing is more complex than that can be accommodated by mora counts: the contour bearing advantage of syllables in prosodic-final position and in words of fewer syllables cross-linguistically cannot be properly represented moraically, and the moraic account has nothing to say about the falling-rising asymmetry and the greater restrictions of more pronounced pitch contours, such as on CVO syllables. To this end, Zhang (2004) proposed that the TBU is the sonorous rime portion of the syllable and that contour tone licensing is gradiently related to the phonetic duration of this TBU. The existence of cases of contour assimilation and dissimilation, either with or without the register feature, has also been argued at length by Bao (1990, 1999) and Yip (1980/1990, 1989, 1995). The formal relation between Register and Contour under a syllabic level TBU rests on whether the register and contour shape of the tone can spread independently in tone sandhi, and arguments on either side can be found in the literature (see Yip 1995 for a review).

From the discussion above, it is clear that the field is far from reaching a consensus on the nature of the TBU and the formal representation of tone. The issue seems stymied by the fact it is always possible to reinterpret a sandhi pattern reported in the literature according to a particular theoretical bent – whether the Danyang /24-T-. . . -T/ → [42-42-. . . -24] sandhi constitutes contour spreading is a case in point (see Bao 1990, 1999; Duanmu 1994; Yip 1989; Chen 2010). I outline two diagnoses for this state of affairs here and then point to directions in which better answers may be sought for the representational issues.
First, the data sources for Chinese tone patterns that phonologists rely on for theorizing are typically impressionistic transcriptions by traditionally trained dialectologists. The transcription of tones is notoriously difficult, and even seasoned fieldworkers may disagree on transcriptions in the same language. For example, the tonal inventory of Tianjin Chinese has been described in three different ways by three different sources, as shown in (8). In Shi’s description, Tones 1 and 2 are considered level tones, while the other two descriptions treat them as contour tones; in Li and Liu’s description, Tone 3 is treated as a complex concave contour, while the other two descriptions treat it as a simple contour. These discrepancies have profound effects on how the tonal patterns are analyzed. For instance, the sandhi Tone 4 → Tone 2 / __ Tone 1 can be interpreted as resulting from a markedness restriction against two falling contours (8b, c) or two adjacent low pitch targets (8a), and the structural change can be interpreted as the metathesis of two Contour features (8b, c) or simply the deletion of one of them (8a). In other words, an arbitrary choice of which source to consult may completely change the analysis and hence the tonal representation that it supports or rejects. This point was also raised in Zhang (2010).

(8) Tianjin tonal inventory:

<table>
<thead>
<tr>
<th></th>
<th>Tone 1</th>
<th>Tone 2</th>
<th>Tone 3</th>
<th>Tone 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Shi (1990)</td>
<td>11</td>
<td>55</td>
<td>24</td>
<td>53</td>
</tr>
<tr>
<td>b. Yang et al. (1999)</td>
<td>21</td>
<td>45</td>
<td>13</td>
<td>53</td>
</tr>
<tr>
<td>c. Li and Liu (1985)</td>
<td>21</td>
<td>45</td>
<td>213</td>
<td>53</td>
</tr>
</tbody>
</table>

Relatedly, the distinction between level tones and contour tones that is crucial to the representation issue may simply be ill-conceived due to this very reason. We have seen that whether a tone is transcribed as level or contour can vary greatly. Acoustically, true level tones with unchanged f0 simply do not exist. Therefore, a representational distinction between a level tone with one pitch specification and a contour tone with two seems unmotivated – every tone has more than one pitch that it needs to hit. One may argue that the level vs. contour distinction may be perceptually based, as small pitch excursions are likely unperceived, but the perception of contours is duration-sensitive (Black 1970; Greenberg and Zee 1979), which makes a binary representational distinction difficult to maintain. Articulatorily, the implementation of contour tones is crucially different from that of complex segments such as [kʰ], or contour segments such as [h] or [s], which either involve two different articulators or two distinct states of one articulator, in that it involves the gradual state change of a single articulator – the vocal folds. This also indicates a close affinity between contour tones and level tones.

Second, as mentioned previously, due to tone’s diachronic malleability, the synchronic tone sandhi patterns are often complex and phonetically arbitrary, and typological generalizations about sandhi patterns invariably have exceptions. This makes the practice of using particular phonetic realizations of sandhi patterns as arguments for tonal representation suspect. The presence of cases that look like contour or register spreading, especially when they are rare, does not necessarily
mean that such theoretical maneuvers are warranted, as the sandhi could be due to lexical or allomorph listing (see Zhang and Lai 2008; Zhang et al. 2009, 2011). Conversely, as in any typological searches, the absence of a pattern is never certain unless the search is exhaustive; this is particularly true for tone patterns where the unexpected is almost expected. The bigger issue here is that typological patterns only constitute indirect evidence for the speakers’ phonological knowledge (Kenstowicz and Kisseberth 1979: ch. 5); there is, therefore, an inherent limitation for using them to argue for the nature of phonological grammar, from representations to the formal relevance of phonetics and opacity. See additional discussion in Zhang (2010), Zhang and Lai (2008, 2010), and Zhang et al. (2009, 2011) on how the issue relates to Chinese tone.

While the phonological discussion of the representational issues are somewhat at a stasis, considerable progress has been made in our understanding of the production, perception, and processing of Chinese tone. In a large body of work on the nature of timing and co-articulation of SC tones in different syllable types and speaking rates, Xu (1994, 1997, 1998, 1999, 2001) showed that: (i) contextual tonal variations are better accounted for by asymptotic approximations of tonal targets that include both static and dynamic targets rather than static targets alone; (ii) when the duration of a syllable carrying a contour tone increases due to a slower speech rate, the dynamic portion of the contour shifts to the later part of the syllable, thus keeping the contour target relatively constant; and (iii), the critical points within a tone, such as the f0 peak for a rise and the f0 valley for a fall, are closely aligned with edges of the syllable. Xu and colleagues hence argued in their Target Approximation model (Xu 2004, 2005; Xu and Wang 2001; for an extension beyond lexical tones, see Prom-on et al. 2009) that the underlying units for contour tones consist of dynamic targets such as [rise] and [fall] with a linear movement specification, and the entire syllable, regardless of its segmental composition, is the TBU. Xu’s view on the unit nature of contour tones is supported by Wan and Jaeger’s (1998) speech error study on Taiwanese Mandarin tones, in which they found that the tonal errors are characterized by a predominance of whole-tone substitution errors, a lack of tone splitting or hybrid tones in word-blend and telescoping errors, and a lack of errors that must be accounted for by register or tone feature spreading. Due to problems with purely phonological forms of the argument outlined above, looking into the phonetic implementation and psycholinguistic processing of tone is likely a more fruitful approach for our understanding of tonal representation and the TBU. Currently, the representation that seems to have the most stock is rather simple: the syllable is the unit around which the implementation of tones is planned and hence the TBU, and the underlying tonal target could be either static or dynamically defined over the entire TBU.

4 The analysis of Chinese tone sandhi patterns

The analysis of Chinese tone sandhi patterns has presented considerable challenges to theoretical phonology in both rule-based and constraint-based frameworks. Earlier rule-based analyses often focused on representational issues of
primitive features and their formal relations as well as the TBU. But aside from the methodological problems mentioned earlier, phonologists who use sandhi patterns to argue for particular tonal representations are often content with accounting for a small corner of the entire sandhi system of the language. Consequently, even if the tonal representation helps the analysis of the sandhi patterns in question, it often does not guarantee the analytical success of the entire sandhi system, nor does it necessarily guarantee that the tone sandhi typology overall will benefit from the representation. During the 1990s, Optimality Theory brought to the forefront the issues of markedness and constraint interaction, while representational issues somewhat receded to the background. The field of Chinese tonology was excited to see whether OT brought new possibilities in accounting for the complex tone sandhi patterns in synchronic phonology. The many gallant attempts (e.g., Chen 2000; Hsiao and Chu 2006; Jiang-King 1996; Lin 2008; Wang 2002; Wee 2004, 2010; Yip 1999, 2004; Zhang 1999), however, were only met with limited enthusiasm in the field. Many scholars remain unconvinced that OT is an appropriate model for complex sandhi processes (e.g., see Bao 2003), and synchronic analysis remains lacking for many sandhi systems.

One of the reasons for the limited success in constraint-based analyses of tone sandhi patterns is the phonetic arbitrariness that permeates present-day sandhi systems, primarily due to diachronic changes. Consequently, if we look for the generalizable markedness constraints that OT relies on for the synchronic analysis of tone sandhi, we tend not to get very far.

However, there are two lines of research on markedness in tone sandhi that are likely fruitful. First, in typological works on Chinese tone and tone sandhi (e.g., Yue-Hashimoto 1987; Chen 1991, 1992, 1996, 2000; Bao 1992; Chang 1992; Jiang-King 1999; Zhang 2002), one markedness generalization does come close to being a true universal: syllables that are rich in duration in their sonorous rhyme portion are better bearers of tonal contrasts. In other words, the following implicational statement seems universally true: if a syllable type with sonorous rhyme duration \( d \) can carry \( n \) tonal contrasts, then a syllable type with sonorous rhyme duration \( d + d_0 \) can carry at least \( n \) tonal contrasts. The syllable-type parameters that may influence the sonorous rhyme duration of a syllable include the syllable's segmental composition, stress property, position in the word (final vs. non-final), and the number of syllables in the word that the syllable belongs to (Zhang 2002). We have seen this generalization in action in the consistently smaller tonal inventories on CVO syllables than CVV/CVR syllables across Chinese dialects. Its manifestation in tone sandhi patterns is also ubiquitous. Right-dominant sandhi systems often involve tonal neutralization in non-final syllables due to their lack of final lengthening effects (Zhang 2002). Both the SC and Taiwanese sandhi patterns in (1a) and (1b) involve the reduction of tonal contrasts by one in non-final position. In the Changzhou sandhi pattern in (3), the rightward tone spreading causes syllables in di- and trisyllabic words to lose the original tone contrasts except for the initial syllable, while monosyllabic words preserve all tonal contrasts.\(^5\) It is worth noting, however, that what the original tonal inventory will neutralize to under short duration often remains unpredictable. In both SC and Taiwanese, the historical basis
for the tone sandhi is clear, but why 213 must neutralize to 35 in SC or why 24 and 55 must neutralize to 33 in Taiwanese is less so, as even if markedness principles can be proposed, they will not be easily generalizable to related dialects that have the same historical sandhi. Therefore, we may need to be content with simply predicting neutralization under adverse durational conditions rather than the exact tones that will surface under sandhi. The smaller inventory may be derived à la Flemming’s (1995/2002, 2003) Dispersion Theory, but the exact tone to surface for a particular sandhi may be the result of allomorph listing – an approach championed by Tsay and Myers (1996), Yip (2004), and a series of works by Zhang and colleagues (Zhang and Lai 2008; Zhang et al. 2009, 2011).

Another line of research is on the psychological reality of phonetically-based phonological markedness. As previously stated, in general, typological studies have only been able to identify soft markedness universals in tone sandhi, while phonetically unexpected patterns are ubiquitous. Does this then mean that no synchronically useful markedness principles can be formulated? This question goes to the heart of the debate on the relevance of phonetics to synchronic phonology. The exceptions rule out the hard-line position that phonetically based constraints, intrinsic rankings, grounding conditions, or other formal mechanisms are hardwired, and some have taken them as evidence that the effect of phonetics on phonological typology only takes place in the realm of diachronic sound change (e.g., Hyman 2001; Yu 2004; Blevins 2006). But it is also possible that the design scheme of the grammar only includes an analytical bias that favors the learning of patterns with stronger phonetic bases (Wilson 2006). This type of approach predicts strong universal tendencies in favor of phonetically motivated patterns, but allows “unnatural” patterns to surface. Chinese tone sandhi provides us with many opportunities to test the analytical bias hypothesis: if the productivity of a sandhi pattern in novel words is correlated with the phonetic nature of the sandhi, in that a pattern with a stronger phonetic basis is more productive than one with a weaker phonetic basis, then despite the exceptions to the markedness generalizations, there is still an analytical bias in favor of the phonetically unmarked pattern.

To conduct this type of study, we need two sandhi patterns that satisfy the following conditions: (i) they have comparable triggering environments; (ii) they are of comparable productivity in the lexicon; (iii) they have comparable lexical frequencies in the lexicon; and (iv) they differ in their degrees of phonetic motivation. Zhang and Lai (2010) investigated the productivity difference between the “third-tone sandhi” (213 → 35 / __ 213) and the “half-third sandhi” (213 → 21 / __ T, T ≠ 213) in SC with promising results. Upon arguing that the “half-third sandhi” has a stronger phonetic basis than the “third-tone sandhi,” Zhang and Lai showed that the former applies more productively to novel words than the latter in two wug-test (Berko 1958) experiments, thus supporting the analytical bias approach to phonetically based markedness.6 There are likely many other Chinese dialects in which we can find sandhi patterns with different degrees of phonetic motivation, and testing the productivity of these sandhi patterns may shed further light on the issue of tonal markedness and the phonetics–phonology interface debate. The likely theoretical consequence is that although many sandhi patterns may still
need to be listed, there may be analytical biases in favor of the less marked ones in the learning process, thus giving them an advantage in typological occurrence and productivity, but in the meantime allowing marked patterns to exist.

Therefore, the analysis of tone sandhi that truly captures the speakers’ phonological knowledge may be dramatically different from an analysis that is solely based on the elicited data pattern. Take Taiwanese tone sandhi, for example: early experimental works by Hsieh (1970, 1975, 1976) and Wang (1993) showed that despite the regularity of the sandhi pattern in the Taiwanese lexicon, the sandhis are largely unproductive when speakers are tested with novel words in a wug test. A series of new experiments by Zhang and colleagues (Zhang and Lai 2008; Zhang et al. 2009, 2011) further quantified the productivity results in a number of morphological contexts and showed that sandhi productivity is not only negatively affected by the opacity of the pattern, but also influenced by the lexical frequency and phonetic nature of the pattern. To this end, Zhang and colleagues proposed stochastic grammars and learning models that can capture both the regular sandhi behavior in the Taiwanese lexicon and the variable sandhi productivity in novel words using extensive lexical listing constraints as well as analytically biased markedness constraints.

The lesson that we learn from these experimental works is that what synchronic phonology needs to account for can be very different from the sandhi patterns observed in the lexicon. It may be the case that there is good reason for the lack of a straightforward synchronic analysis for a pattern, as the pattern is simply unproductive. Even for those lexical patterns for which an analysis seems readily available, the question of whether they are a true reflection of the speakers’ knowledge is still relevant. In other words, to address the analytical conundrum of tone sandhi patterns, we may need to go back and ask a more basic question about the empirical data: Which parts of the patterns are productive, and which parts are not?

5 Variation, gradience, and exceptions in Chinese tone patterns

In a recent trend, phonological research is paying increasingly close attention to the roles of variation, gradience, and exceptions in the grammar. Observations that many phonological patterns are variable (e.g., Labov 1972, 1994), gradient (e.g., Bolinger 1961), and full of exceptions (e.g., Zimmer 1969) have been longstanding, but only recently did experimental works start to show that the observed variability, gradience, and patterns of exceptionality are reflected in the speakers’ phonological knowledge and warrant formal phonological analyses (e.g., Frisch and Zawaydeh 2001; Zuraw 2007; Hayes et al. 2009), and variations of Optimality Theory such as Stochastic OT (Boersma and Hayes 2001), Harmonic Grammar (Smolensky and Legendre 2006), and Maximum Entropy Grammar (Goldwater...
and Johnson 2003; Jäger 2007) have been applied to some of the experimental results in order to capture them formally.

The study of Chinese tone is well situated in the investigation of variation, gradience, and exceptions in phonological theory. As in other languages, the reasons for the presence of these properties in Chinese are complex, but the following characteristics of modern Chinese dialects particularly stand out as factors that encourage variation and exceptions in their tonal patterns. First, different Chinese dialects may have very different tonal inventories and tone sandhi systems, but the vast improvement in transportation and media outreach, as well as the mobility of a large migrant worker population, have brought different dialects considerably closer to each other than in the past. The increased dialectal contact has caused greater instability of the phonological systems and created a fertile ground for variation and exceptions to appear in these systems. Second, with the promulgation of SC as the standard language in Mainland China, the phonological system of SC has had a strong influence on many Chinese dialects, especially Mandarin dialects that are similar to SC to begin with. Third, the difference between written and spoken Chinese has had a long history in China, and many dialects have lexical items with different pronunciations depending on the colloquial or literary style, causing variation.

Chinese linguists have long noted the effects of dialectal contact and the influence of a dominant standard language, and descriptive works that outline dialectal changes due to these two factors abound. For example, the changes within the Shanghai dialect from older to newer generations as a result of influence from other dialects were extensively documented in Xu and Tang (1988), Zhu (2006), You (2006a, b), among others. Based on acoustic results, Shi and Wang (2004) and Zhang and Liu (2011) discussed the variable changes in both the shapes of the lexical tones and tone sandhi patterns in Tianjin and how the changes may have derived from the dialect’s close interaction with SC. The implications of these data for phonological theory, however, remain largely unexplored. It would be interesting to explore what the synchronic grammar looks like to allow the variable patterns to surface, and whether there is a model of the speaker that can predict the effects of contact on the phonological systems of the dialects in question.

The variable pronunciation based on style has also been a steady focus of research by Chinese linguists. Descriptive works for individual dialects and theoretical works both exist, but the theoretical works primarily have a historical focus (e.g., Liu 2003; Wang, H.-J. 2006; Wang, F.-T. 2009). It would be interesting to explore how the synchronic speakers have internalized the literary and colloquial lexical strata and constructed a grammatical model accordingly.

The points made in this section echo the earlier point that phonological research on Chinese tone can significantly benefit from a rebuilt empirical basis: variation, gradience, and exceptions are par for the course for phonological patterns, and tonal patterns in Chinese are no exception. The variable and gradient nature of the patterns that empirical research reveals for Chinese tone presents an excellent opportunity for its contributions to current phonological theory.
6 Other issues

6.1 Structure-sensitive tone sandhi

The tone sandhi behavior in Chinese dialects is often sensitive to the morpho-
syntactic structure of the word. For example, in Shanghai, modifier-noun disyllabic
words have a left-dominant sandhi pattern that involves tone spreading from
the first syllable; but verb-object disyllabic words, especially ones with low lexical
frequencies, often have a right-dominant sandhi pattern that involves contour
reduction on the initial syllable (Xu and Tang 1988; Zhu 2006). In Pingyao, words
with a subject-predicate or verb-object structure also have a different set of tone
sandhi behavior from words with other structures, such as modifier-noun (Hou
1980). Whether the sandhi difference is morpho-syntactically based or prosodically
based is controversial, and the productivity of these sandhi patterns in novel
contexts is largely unknown. The investigation of this issue will shed light on the
interface between phonology and morpho-syntax. For more detailed discussion
of this issue, see Simpson, Chapter 18, this volume.

6.2 The relation between tone sandhi and stress

Related to the morpho-syntactic dependency in tone sandhi, is the relation between
tone sandhi and stress. Duanmu (2007) argued that compounds in SC have gram-
matical stress determined by the “Non-head Stress” principle of Cinque (1993)
(see Chapter 18); for example, a disyllabic modifier-noun compound has stress on
the initial syllable, while a disyllabic verb-object phrase has stress on the final
syllable. Despite the fact that such stress has no clear acoustic correlates in dura-
tion or pitch due to the tonal nature of the language, Duanmu showed that it must
exist based on generalizations about the length and ordering of different gram-
matical components in compounds and phrases. Duamnu (1995) applied similar
arguments to Shanghai and Taiwanese and showed that modifier-noun com-
pounds have initial stress in Shanghai, but final stress in Taiwanese. In dialects
with structure-sensitive tone sandhi, the dominant edge of the sandhi is often
correlated with the location of grammatical stress, such as the Shanghai pattern
mentioned above. It remains to be seen whether grammatical stress similarly
established in other dialects of Chinese also lacks acoustic correlates typically
associated with stress (but see Zhu 1999 for Shanghai), but in any event, the
relevance of this prosodic factor complicates the search for the cause of structure-
sensitive tone sandhi.

Stress that is independent of morpho-syntactic structure has also been reported
in Chinese dialects, and its relation to tone sandhi has been an uncomfortable one
in the study of Chinese tone. For example, Rose (1990) reported that in the Northern
Wu dialect Zhenhai, disyllabic words with MH or H on the first syllable have
initial stress, while those with ML or L on the first syllable have final stress; this
stress pattern was confirmed by native speaker perception. The tone sandhi
pattern in Zhenhai, however, is not easily predicted by this stress pattern, and Li
(2003, 2005) showed that a good understanding of the pattern can only be achieved when we carefully tease apart the interaction between this tonally induced stress and initial prominence, which provides the initial syllable with a longer duration. In the Southern Wu dialect, Wenzhou, both an experienced fieldworker (Zhengzhang 1964) and a native speaker (Cao 2003) reported initial stress, yet the tone sandhi pattern of the language is largely right-dominant with tonal neutralization on the initial syllable. In Danyang, a Northern Wu dialect, stress is initial, as reported by Lü (1980), yet a subset of the sandhi patterns is right-dominant, as argued by Chan (1991). Chan (1995) subsequently offered a diachronic explanation for the mismatch between stress and sandhi direction: the right-dominant sandhi pattern is a vestige of a historical stage of the language with final stress.

What is often lacking in these sources is again careful phonetic descriptions of the acoustic correlates for stress and tone like Rose did for Zhenhai. A comparison of productivity between sandhi patterns that agree with stress assignments and those that do not will likely also be enlightening. It is hoped that empirical studies will go hand-in-hand with theoretical advances to shed light on the relation between stress and tone, especially in languages where they seem to conflict with each other.7

6.3 Tone sandhi in longer sequences

Tone sandhi often applies to sequences longer than two syllables. This is well documented in many Chinese dialects, for example SC (Shih 1986), Taiwanese (Chen 1987; Lin 1994), Pingyao (Hou 1982a, b), Shanghai (Xu and Tang 1988; Zhu 2006), and Tianjin (Li and Liu 1985). The sandhi behavior may be sensitive to the morpho-syntactic structure of the sequence (SC, Taiwanese, Pingyao, Shanghai), but may not be (Tianjin). The analysis of this type of sandhi has proven difficult. For instance, the directional sensitivity of sandhi application in long sequences in Tianjin Chinese and the difficulties it presents to Optimality Theory feature prominently in the literature on Tianjin tone sandhi (e.g., Chen 2000; Lin 2008; Wee 2010); the Changting Hakka trisyllabic tone sandhi pattern has been argued by Chen (2004) to be unanalyzable by current theoretical apparatuses available to an OT phonologist. But the role that this type of sandhi plays in our understanding of tone sandhi domain, the correspondence between morphologically related forms, and the structure of the grammar as either derivational or parallel, is undeniable. Many of the longer sequences are phrases. Therefore, the productivity of the sandhi pattern is less in doubt. But the field still has much to gain from careful phonetic and psycholinguistic testing of the patterns without purely relying on impressionistic transcriptions of the original fieldworkers.

7 Conclusion

This chapter has reviewed some past and current issues in the tonal phonology of Chinese in this chapter. Typologically, Chinese tone systems are characterized
by complex tonal inventories and tone sandhi patterns, and the sandhi patterns are often phonetically arbitrary. These properties present both opportunities and challenges for our understanding of tonal representations and tonal analysis. This chapter has argued that phonological representations and analyses of Chinese tone based on traditional descriptions of lexical patterns run the risk of being premature, as the speakers’ knowledge of tonal patterns may not be identical to the patterns in the lexicon and impressionistic transcriptions, no matter how careful, have their limitations. The study of Chinese tones, therefore, has much to gain from carefully designed phonetic and psycholinguistic investigations, and such investigations have begun to shed light on the nature of tones and tone sandhi. Finally, the fact that Chinese tone patterns are rich in variation, gradience, and exceptions should continue to capture the attention of phonologists working in this area, and models of tonal grammar should aim to capture not only categorical regularities, but also the variable and gradient part of the speakers’ knowledge.

NOTES

1 Tones here are transcribed in Chao numbers (Chao 1948, 1968), where “5” and “1” indicate the highest and lowest pitches in the speaker’s pitch range, respectively. Juxtaposed numbers represent contour tones; for example, “51” indicates a falling tone from the highest pitch to the lowest pitch.

2 The Jin dialects spoken in Shanxi province, sometimes considered as part of the Northern family, have kept the glottal coda [Ɂ]. This also prompted some researchers to consider them as a separate dialect group (Hou and Wen 1993). The tonal inventories on obstruent-closed syllables in these dialects behave similarly to those in the other dialect groups.

3 Two prominent Chinese journals – Fangyan (Dialects) and Zhongguo Yuwen (Studies of the Chinese Language) – are particularly instrumental in the publication of these descriptions.

4 Other types of tonal representations can also be found in the literature. Barrie (2007), for example, assumes Dresher’s (2003) Theory of Contrastive Hierarchy and represents tones in the contrastive hierarchy Register > Pitch > Contour. Contour tones are unitary segments with the feature [contour] under this analysis.

5 Even this markedness generalization may arguably have exceptions. There are occasional Chinese dialects whose full tonal inventory is only realized in sandhi forms, but partly neutralized in monosyllabic isolation forms. For example, in Pingyao (Hou 1980), tone 13 has two different sets of sandhi behavior when appearing on the initial syllable of disyllabic words with certain grammatical structures. Essentially, the analysis needs to recognize that two different underlying tones are neutralized in monosyllabic forms, but surface distinctly in disyllabic forms. However, Pingyao is not a typical right-dominant language – the second syllable of disyllabic words also undergoes sandhi in Pingyao. Therefore, the realization of the tonal contrast on the first syllable is not confined to the duration of the first syllable. In fact, only a limited number of di-tonal combinations occur in disyllables, indicating neutralization of tonal patterns found in disyllable. Whether there exist true exceptions to the durationally based markedness generalization requires further research.
Zhang and Lai (2010) also ruled out alternative approaches, such as lexical frequency, priming, and the nature of lexical listing, as being solely responsible for the results.

There is another type of stress in Chinese: many Chinese dialects also have a type of lexically specified stress-less syllables, referred to as “qing sheng” (“light sound”). For example, in Beijing Chinese, *qing sheng* can appear with both grammatical morphemes (*diou55 le0 “lose-perfective”*) and lexical morphemes (*cong55 ming0 “smart”*). These stress-less syllables have a drastically reduced duration and, correspondingly, drastically reduced tonal inventories. Beijing Chinese, for example, has no contrastive tone on this type of syllable. There does not seem to be a mismatch between this type of stress, or rather, stress-lessness, and tone. This is another indication that a good understanding of tonal behaviors needs to come from empirical phonetic data beyond pure transcription labels such as [-stress] (see also Liu 2010).

**REFERENCES**


Wang, H. 2006. Wenbai yidu, yinyun cengci yu lishi yuyanxue [Literary vs. colloquial pronunciations, phonological strata, and historical linguistics]. *Beijing Daxue Xuebao – Zhexue Shehui Kexue Ban*


Yip, M. 2004. Phonological markedness and allomorph selection in


