Chapter 12

A Two-level Take on Tianjin Tone

Martin Jansche

Abstract. The tonal alternations of Tianjin Mandarin are investigated. The data, which have been problematic for rule-based approaches, receive an analysis in a declarative two-level setting. Formal differences between certain tonal alternations are conjectured to bear on speech production but not on acquisition.

1 Introduction

The regular tonal alternations exhibited by the dialect of Mandarin spoken in the city of Tianjin, P. R. China, have received quite a bit of attention since Matthew Chen first pointed out that they, in his words, "appear to severely strain the descriptive and explanatory capability of current linguistic theory" and "call into question just about every assumption linguists have long held about the mode of application of phonological rules and their interaction with [...] prosodic organization" [Chen, 1986, pp. 98, 111]. "The paradox of Tianjin tone sandhi"—featured as the title of Chen's original talk—arises when one tries to determine the mode of application of a small inventory of seemingly straightforward rules meant to formally capture the alternations. I will argue against the traditional accounts of the Tianjin tone sandhi not only because the proposed analyses are often factually inadequate, but also because they obscure issues relating to the processing and acquisition of these tonal alternations. The inherent complexity of two alternations invites speculation about their difficulty for incremental speech production.

This paper is organized as follows. After a brief overview of the facts and issues (S2), I will point out why the Tianjin tone sandhi have resisted a satisfactory analysis within the frameworks of mainstream phonology (S3). These accounts contrasted with the elegance of a two-level, finite-state account of the data, which also allows me to address the relative complexity

Proceedings of the Third ESSLLI Student Session
Ivana Kruijff-Korbayová (editor).

of the tonal alternations (S4). I conclude with a discussion of the implications for production, acquisition and automatic induction, and diachronic explanations for the present-day facts (S5).

2 The facts

The Tianjin tone sandhi have become the focus of quite a few recent studies [Chen, 1986, Chen, 1987, Hung, 1987a, Hung, 1987b, Tan, 1987, Zhang, 1987, Chen, 1995a] since they were described in [Li & Liu, 1985], and have recently acquired the status of a textbook example (cf. [Kenstowicz, 1994, pp. 328f.]; [Yip, 1995, p. 491]).

Like many other Northern Mandarin dialects, Tianjin Mandarin has four tone(me)s. In isolation, the four citation tones are: first, low falling [21]; second, high rising [45]; third, low falling-rising [213]; and fourth, high falling [53] [Chen, 1986, p. 98]. The numbers indicate pitch on an idealized five-point scale, with '5' denoting the highest pitch. The exact tonal shapes are not uncontroversial, but fortunately don't play a crucial role for the remaining discussion. I will simply stick with Chen's numbers for ease of comparison. The first tone and the third tone each have one, the fourth tone two sandhi tones, all of which are taken from the paradigm of the isolation tones (this is called *paradigmatic replacement* in [Wang, 1967, p. 94]), i.e., the range of possible surface forms is identical to the set of tones found in isolated syllables.

My data are mainly taken from [Chen, 1986], which in turn is based on [Li & Liu, 1985], partly superseded by additional data from [Tan, 1987], with some qualifications I felt compelled to make in the light of the discussion set forth there. Where the literature is in disagreement as to what the exact facts are, I take the more complex patterns as basic, from which the simpler ones can be explained by different prosodic phrasing. Since the segmental details are not at stake, I have taken the liberty of using *Hànyǔ pīnyīn* throughout instead of IPA transcriptions (cf. [Chen, 1987, n. 5] and [Hung, 1987a, ch. 6, n. 4]). The numbers after the glosses serve to locate the examples in [Chen, 1986]. The glosses have been altered slightly.

2.1 First tone sandhi

The first tone sandhi changes underlying /21/ into a surface third tone [213] when it immediately precedes [21].

(12.1)	gao shan	'high mountain' (C86–2.1)
	21 21	underlying tones
	213 21	surface tones

That the triggering environment is a surface tone [21] is evident from (12.2), where the [21] sandhi tone of /53/ triggers the change of /21/ to [213].

(12.2)	tong [dian huar]	'make a phone-call' (C86–4.1)
	21 53 53	underlying tones
	213 21 53	surface tones

That the triggering environment is not an underlying tone /21/ can be seen from (12.3), in which the leftmost surface tone is identical to the underlying tone (also note that the sandhi is not directly sensitive to morphological structure). When several underlying tones /21/ appear in a row as in (12.4), the surface form shows a characteristic alternation between [213] and [21] with the rightmost tone surfacing as [21].

(12.3)	kai [fei ji]	'fly an airplane' (C86–3.1)
	[gao ya] guo	'pressure cooker' (C86–11)
	21 21 21	underlying tones
	21 213 21	surface tones

(12.4)	kai [[tuo la] ji]	'operate a tractor' (C86–23)
	21 21 21 21	underlying tones
	213 21 213 21	surface tones

2.2 Third tone sandhi

The third tone sandhi changes an underlying third tone /213/ into a surface second tone [45] when it immediately precedes '213'.

(12.5)	xi lian	'wash one's face' (C86–2.2)
	213 213	underlying tones
	45 213	surface tones

The third tone sandhi can either be triggered by a surface third tone [213] as in (12.6), where [213] is the sandhi tone of /21/, or by an underlying third tone /213/ as in (12.7). (This opaque interaction has been noted in passing by [Chen, 1986, p. 109], but was not pursued any further by him.) Multiple underlying third tones exhibit a surface pattern in which all tones but the last one surface as the sandhi tone [45].

(12.6)	ting	[guan	xin]	'quite concerned' (C86–4.2)
	[bao	wen]	bei	'thermos cup' (C86–8)
	213	21	21	underlying tones
	45	213	21	surface tones

(12.7)	chang	[dang	wei]	'factory party commissar' (C86–13.1)
	[li	fa]	suo	'barbershop' (C86–12.a)
	[ma	zu	ka]	'mazurka' (loanword) (C86–16.b)
	213	213	213	underlying tones
	45	45	213	surface tones

2.3 Fourth tone sandhi

The fourth tone sandhi changes an underlying /53/ into [45] before [21]:

(12.8)	kan	shu	'read books' (C86–2..4)
	53	21	underlying tones
	45	21	surface tones

Note that the triggering environment is only the surface tone [21], since no sandhi takes place when /53/ precedes an underlying tone /21/ with a surface tone in sandhi form, as seen in (12.9). On the other hand, [21] still triggers a sandhi tone for /53/ to its left, even if [21] itself is a sandhi tone, which is illustrated by (12.10).

(12.9)	ji	[gong	fen]	'record work points' (C86–3.2)
	[zi	zun]	xin	'self-esteem' (C86–16.c-3)
	53	21	21	underlying tones
	53	213	21	surface tones

(12.10)	du	[mi	yue]	'spend one's honeymoon' (C86–14)
	[liu	wan]	si	'64 000' (C86–16.c-1)
	[yi	da	li]	'Italy' (loanword) (C86–16.c-1)
	53	53	53	underlying tones
	45	21	53	surface tones reported in [Tan, 1987]

Note that Chen gives [53 21 53] for the surface tones here. The status of the leftmost surface tone is somewhat unclear, and sometimes the alternation is declared to be optional [Hung, 1987b, p. 286]. However, [Tan, 1987, pp. 230f.] argues that tone sandhi rules apply unconditionally within one tone group. If any one alternation appears to be optional, the string has been broken up into smaller tone groups, resulting in a reduced number of changes. I will assume a worst-case scenario here, in which the change of /53/ to [45] takes place obligatorily within one tone group. I replace the surface tones in Chen's original example, made up of two tone groups [53 | 21 53], by the single tone group [45 21 53] (cf. [Tan, 1987, p. 234, ex. 7; p. 241, ex. 22; p. 243, ex. 26]).

The last example already illustrated the second sandhi form of the fourth tone /53/: unless it precedes [21], it changes into [21] before /53/.

(12.11)	jing	zhong	'net weight' (C86–2.3)
	53	53	underlying tones
	21	53	surface tones

Note that this change is triggered by an underlying fourth tone /53/ only, unless that happens to have a surface form [21].

(12.12)	zuo	[dian	che]	'take a tram' (C86–13.2)
	[dian	shi]	ji	'television set' (C86–16.c-2)
	53	53	21	underlying tones
	21	45	21	surface tones

In sum, an underlying fourth tone /53/ will surface as [45] if it precedes [21], or otherwise as [21] if it precedes /53/. Thus the fourth tone sandhi exhibits the kind of opacity also found in the third tone sandhi, since the change that produces [21] is sensitive to the following underlying tone, whereas the change that produces [45] is similar to the first tone sandhi in that it is only sensitive to the surface form of the following tone.

3 An apparent paradox

The seemingly paradoxical aspect of the Tianjin tone sandhi according to [Chen, 1986] is the lack of a principled and uniform mode of application for the innocent-looking rules he proposes. His discussion is reminiscent of [Yip, 1980, S5.1], which concerned the modes of rule application in the case of the Beijing Mandarin third tone sandhi, but Chen reaches the conclusion that no single mode of application can derive all polysyllabic tonal alternations of Tianjin Mandarin.

The following presentation of this "paradox" summarizes the discussion in [Zhang, 1987, S2].

At first sight it doesn't seem too implausible to postulate the sandhi rules given in (12.13).

(12.13) Tianjin tone sandhi rules [Chen, 1986, p. 98]

TS1	21	⟶	213	/ __	21
TS3	213	⟶	45	/ __	213
TS4a	53	⟶	21	/ __	53
TS4b	53	⟶	45	/ __	21

Given the conventional understanding of these rules, the bisyllabic patterns found in (12.1), (12.5), (12.11), and (12.8) can be derived without problems. But for longer sequences, no single mode of rule application will work uniformly in all instances.

(12.14) cyclic rule application

(a) works in some cases

kai	[fei	ji]	'fly an airplane'
21	21	21	underlying tones
	213		TS1 on inner cycle
n/a			TS1 not applicable
21	213	21	actual surface tones

(b) but not always

[bao	wen]	bei	'thermos cup'
213	21	21	underlying tones
n/a			no rule applicable on inner cycle
	213		TS1 on outer cycle
*213	213	21	predicted surface tones
45	213	21	actual surface tones, would require post-cyclical application of TS3

(12.15) simultaneous/parallel/across-the-board rule application

(a) works for some of the examples some of the time

chang	[dang	wei]	'factory party commissar'
213	213	213	underlying tones
45	45	213	TS3 produces actual surface tones

(b) but not for all of the examples all of the time

du	[mi	yue]	'spend one's honeymoon'
53	53	53	underlying tones
*21	21	53	TS4a produces incorrect results
45	21	53	actual surface tones

(12.16) directional rule application

(a) sometimes works right to left

[gao ya] guo	'pressure cooker'
21 21 21	underlying tones
213	TS1 applies to rightmost pair
n/a	TS1 not applicable to leftmost pair
21 213 21	actual surface tones

(b) sometimes doesn't work right to left

[li fa] suo	'barbershop'
213 213 213	underlying tones
45	TS3 applies to rightmost pair
n/a	TS3 not applicable to leftmost pair
*213 45 213	predicted surface tones

(c) but rather left to right

[li fa] suo	'barbershop'
213 213 213	underlying tones
45	TS3 applies to leftmost pair
45	TS3 applies to rightmost pair
45 45 213	actual surface tones

(d) left to right application of TS1 in (a) would of course fail miserably

Subsequent analyses [Hung, 1987a, Hung, 1987b, Tan, 1987, Zhang, 1987] have used rule-specific directionality stipulations, extrinsic rule-ordering, constraints on surface forms, etc. [Chen, 1987] provides an assessment of various shortcomings of these proposals, none of which appears to be completely satisfactory. Chen's own recent attempts [Chen, 1995b, Chen, 1995a] to analyze the Tianjin data in terms of "derivational optimality" do not seem to fare any better, because of the unnecessary conceptual overhead. In [Chen, 1995b] he briefly mentions a two-level "optimality" approach, but immediately rejects it mainly because this kind of solution appears too stipulatory and arbitrary to him. In the remainder of this paper, I argue that Chen was too quick to dismiss this approach, since it can be put to use by stripping away all "optimality" machinery. The resulting declarative two-level analysis is not only descriptively adequate and thus superior to many preceding attempts while being conceptually simpler, but also gives us an opportunity to integrate the phonological description with theories of processing and acquisition, which all other approaches fail to address.

4 Paradox lost

The correspondences between underlying tones and surface tones observed in S2 can be expressed very elegantly as constraints within the framework of Finite-State Phonology/Morphology ([Karttunen, 1993] is an excellent overview), or equivalently as finite-state transducers. The particular finite-state constraints that derive the Tianjin sandhi can easily be read off the descriptive generalizations arrived at in S2. The corresponding transducers let us focus on some computational aspects that cannot easily be gleaned by looking at the finite-state constraints, which is why one representative transducer deserves some further discussion.

The observed tonal alternations within one tone group are described as a rational function in the set $\{21, 213, 45, 53\}^*$, as defined by the following finite-state transducer with a set of states $Q = \{0, 1, 2, 3, 4, 5, 6\}$, an input and output alphabet $\Sigma_{\text{in}} = \Sigma_{\text{out}} = \{21, 213, 45, 53\}$, an initial state $0 \in Q$, a set of final states $\{0, 1, 3, 5\} \subset Q$, and a partial transition function $Q \times \Sigma_{\text{in}} \times \Sigma_{\text{out}} \to Q$ corresponding to the following transition table:

(12.17)

	21	21	45	213	213	53	53	53
	21	213	45	213	45	53	45	21
0.	1	2	0	3	4	5	2	6
1.		2	0	3	4	5	2	
2:	1							6
3.	1		0			5	2	6
4:		2		3	4			
5.		2	0	3	4			
6:						5	2	

As an illustration of how the machine works as an acceptor, see what happens when it accepts the string of pairs '21:21 21:213 21:21' by switching its internal state from the start state 0 to 1, from 1 to 2, and from there to 1, one of the designated final states. In fact when both the input and output are known, it always behaves like a deterministic automaton. Note that although its underlying finite-state automaton is deterministic, the transducer itself is inherently nondeterministic, i. e., it cannot be transformed into an equivalent deterministic transducer (cf. [Savitch, 1982, pp. 69f.]).

To see this, consider the behavior of the machine when it is operating as a transducer proper. For instance, it would map an input of /21 21 21/ to the unique output [21 213 21], or an input string /21 21/ to the corresponding output string [213 21]. Here the first token in the output string corresponding to the input with a length of 2 is [213], whereas it is [21] if the input string is 3 tokens long. This is true for all inputs of the form $/21^+/$ of even or odd length, respectively, as can be shown by induction on the length of the input. This incidentally also shows that the transducer is lacking the prefix property that is needed in order for it to be sequential (cf. [Mohri, 1997, p. 360]). If the transducer is given a string of the form $/21^{+}/$ as input, it has to choose its first output token nondeterministically. To be sure that this choice was correct, the transducer would have to scan the whole input string and count the number of tokens. As a corollary, it follows that the transducer wouldn't always produce an output if we allowed infinitely long input strings.

Now, humans hardly ever try to produce infinitely long outputs. Still, the above insight is not totally meaningless if we assume that humans have limited memory resources and are not capable of massively parallel processing for incremental speech production. This might explain why no reasonably long strings of $[(213\ 21)^n]$ output corresponding to underlying $/21^{2n}/$ have been documented in the literature. Humans either make mistakes in this case, or break up the input into smaller, more easily processable parts by adding additional boundaries across which the sandhi processes cannot operate.

A similar statement could be made about the fourth tone sandhi. But what is really important is the way in which both differ from the third tone sandhi. To decide whether an input of /213/ comes out as [213] or [45], the transducer only has to wait for the next input token. If it is /213/ again, the output corresponding to the previous token must be [45]. Ignoring the fact that the first and third tone sandhi interact, we can say that the former is not decidable with any fixed lookahead, while the latter one is decidable with a lookahead of 1. We can easily specify a sequential transducer that implements the third tone sandhi: take a set of states $Q = \{0, 1, 2\}$ with 0 the initial state and $\{0, 2\}$ the final states, and leave the alphabets implicit in the definition of the partial transition (and output) function $Q \times (\Sigma_{\text{in}} \cup \{\$\}) \rightarrow Q \times \Sigma^*_{\text{out}}$ ($\varepsilon \in \Sigma^*_{\text{out}}$ is the empty string, '\$' marks the end of the tone group):

(12.18)

	45	213	\$
0.	$\langle 0, 45 \rangle$	$\langle 1, \varepsilon \rangle$	
1:	$\langle 0, 213\ 45 \rangle$	$\langle 1, 45 \rangle$	$\langle 2, 213 \rangle$
2.			

It is important to note that the first and fourth tone alternations are different from similar anticipatory phenomena involving leftward spreading, as for example in vowel harmony. In the case of vowel harmony, the vocalic makeup of the root might determine the vowel qualities in a prefix where the distance between the prefix and the root is in principle unbounded. However since vowel quality is highly salient in languages with vowel harmony, one might suspect that speakers have access to this information at an early planning stage during production. But note that in the case of the Tianjin first tone sandhi it is not only the tone of the last syllable in a tone group, which could in principle be an unbounded distance away to the right, but in fact the exact length of the tone group that determines the tone of the first syllable of the group in the worst case. It is the required measuring of this distance that leads me to suspect that the first and fourth tone sandhi might prove difficult for incremental speech production.

A second important distinction between the third tone sandhi and its first tone and fourth tone counterparts is in the amount of ambiguity created. A surface string of the form [45^n 213] corresponds to $n+1$ different underlying strings made up from tones /45/ and /213/, possibly involving the third tone sandhi. But where the two other sandhi processes might have applied, the number of underlying forms from which a surface string could have been generated is of a different order: for a surface string of the form [$(213\ 21)^n$] there are 2^n underlying strings (due to the alternating pattern of the first tone sandhi), and a surface string [$(45\ 21)^n$ 45 21 53] corresponds to $4 \cdot 3^n$ underlying strings. Still, I do not expect significant difficulty for human online processing, because of lexical gaps (not all possible underlying strings represent meaningful words or phrases) and, more importantly, because of disambiguation due to the preceding context.

5 Directions for further research

The formal properties of the Tianjin tone sandhi transducer raise a number of questions about how the grammatical system it represents can be learned and used by both humans and machines. Since the transducer is not sequential, familiar techniques for efficient induction cannot be applied. How the above transducer can be learned efficiently (and ideally from positive examples only) is a question that merits further research. It is clear, however, that the differences between the individual tonal alternations do not make a difference for induction. Recall that the nondeterminism of the transduction mode vanishes when one views the transducer as an automaton accepting strings of input/output pairs. The induction task comes down to inferring a suitable transducer from pairings of representative input and output strings, and for this the facts concerning the behavior of the machine in transduction mode are irrelevant. Since the first and fourth tone sandhi are, for the purpose of acquisition, not different from the third tone sandhi, which is very stable across dialects of Mandarin in general (see below), I expect acquisition of these alternations to be no more difficult than acquiring the seemingly less complex third tone sandhi.

Machine processing of the transducer is largely unproblematic: for example, the time for generating a surface form from a lexical representation is linear in the length of the input, given reasonable working memory; recognition is harder, but this is due to the inherent ambiguity of the surface forms noted above. As mentioned earlier, the first and fourth tone sandhi might however turn out to complicate incremental generation for humans. Since in the worst case a lookahead equal to the total length of the input string is required, we would expect humans to break up large units that would trigger the problematic alternations into smaller parts, perhaps producing smaller sandhi domains. Since the central function of the tone sandhi is to indicate the presence of phrase boundaries when a sandhi has failed to apply, we are often able to see from a surface string (given its meaning) if a potentially larger domain has been split into several smaller ones. I would predict that this is to be encountered more frequently with configurations triggering the first and fourth tone sandhi, than (everything else being equal) with the third tone sandhi, which requires a fixed amount of memory. Thus the third tone sandhi provides some kind of control element against which one could test the first and fourth tone sandhi. I leave it for future research to design and carry out an experiment that would verify or falsify these predictions.

Looking for an explanation for the various tonal alternations, one should notice that at least the third tone sandhi of Tianjin Mandarin is not an isolated phenomenon. Many Mandarin dialects exhibit similar phenomena, and both dialectal comparison and external evidence suggest that this alternation was already present in Early Mandarin of the 16th century [Mei, 1977]. It might even be appropriate to speak of a Pan-Mandarin Third Tone Sandhi (cf. [Zhang, 1987, p. 267] and [Hung, 1987b, n. 11]), since this inherited alternation is even at work in dialects where the phonetic shapes of the relevant third and second tones underwent considerable changes. [Bao, 1992, p. 4] gives an example from Luoyang Mandarin, where /yang^{53} ma^{53}/ 'raise horses' surfaces as [yang^{31} ma^{53}]. However, his explanation of this alternation as the result of register dissimilation seems *post hoc*: again all that happens is a third tone /53/ (in the Luoyang dialect) surfacing as a second tone [31], which points to a corresponding regular alternation in an earlier form of the language, at a stage where it perhaps made (more) phonetic sense.

Acknowledgments

Thanks to Mike Calcagno, Janice Fon, Erhard Hinrichs, Nathan Vaillette, and two anonymous reviewers for helpful comments on earlier drafts. The usual disclaimers apply.

Bibliography

[Bao, 1992] Zhiming Bao, 'Toward a typology of tone sandhi.' *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society, February 14–17, 1992: Special Session on the Typology of Tone Languages*, ed. Laura A. Buszard-Welcher, Jonathan Evans, David Peterson, Lionel Wee, and William Weigel. Berkeley Linguistics Society, 1992. 1–12

[Chen, 1986] Matthew Y. Chen, 'The paradox of Tianjin tone sandhi.' *Papers from the General Session at the Twenty-Second Regional Meeting*, ed. Anne M. Farley, Peter T. Farley, and Karl-Erik McCullough. Chicago Linguistic Society, 1986. 98–114

[Chen, 1987] Matthew Y. Chen, 'A symposium on Tianjin tone sandhi: Introductory remarks.' *Journal of Chinese Linguistics* 15.2 (1987): 203–227

[Chen, 1995a] Matthew Y. Chen, 'Directionality and interacting sandhi processes, I.' Ms., University of California, San Diego, December 1995

[Chen, 1995b] Matthew Y. Chen, 'Directionality: Constraints on derivation?' Handout for talk presented at Tilburg Conference on Derivational Residue, October 5–7, 1995. Rutgers Optimality Archive #93d

[Hung, 1987a] Tony T. N. Hung, 'Syntactic and semantic aspects of Chinese tone sandhi.' Ph. D. thesis, University of California, San Diego, 1987. Distributed by Indiana University Linguistics Club, Bloomington, IN, December 1989

[Hung, 1987b] Tony T. N. Hung, 'Tianjin tone sandhi: Towards a unified approach.' *Journal of Chinese Linguistics* 15.2 (1987): 274–305

[Karttunen, 1993] Lauri Karttunen, 'Finite-state constraints.' *The Last Phonological Rule*, ed. John Goldsmith. University of Chicago Press, 1993. 173–194

[Kenstowicz, 1994] Michael Kenstowicz, *Phonology in Generative Grammar*. Blackwell, 1994

[Li & Liu, 1985] Lǐ Xíngjiàn, and Liú Sīxùn, 'Tiānjīn fāngyán de liándú biàndiào.' *Zhōngguó Yǔwén* 184 (January 1985): 76–80

[Mei, 1977] Tsu-lin Mei, 'Tones and tone sandhi in 16th century Mandarin.' *Journal of Chinese Linguistics* 5.2 (1977): 237–260

[Mohri, 1997] Mehryar Mohri, 'On the use of sequential transducers in natural language processing.' *Finite-State Language Processing*, ed. Emmanuel Roche and Yves Schabes. MIT Press, 1997. 355–382

[Savitch, 1982] Walter J. Savitch, *Abstract Machines and Grammars*. Little, Brown and Co., 1982

[Tan, 1987] Fu Tan, 'Tone sandhi in the Tianjin dialect.' *Journal of Chinese Linguistics* 15.2 (1987): 228–246

[Wang, 1967] William S.-Y. Wang, 'Phonological features of tone.' *International Journal of American Linguistics* 33.2 (1967): 93–105

[Yip, 1980] Moira Yip, 'The tonal phonology of Chinese.' Ph. D. thesis, MIT, 1980. Distributed by Indiana University Linguistics Club, Bloomington, IN

[Yip, 1995] Moira Yip, 'Tone in East Asian languages.' *The Handbook of Phonological Theory*, ed. John A. Goldsmith. Blackwell, 1995. 476–494

[Zhang, 1987] Zheng-sheng Zhang, 'The paradox of Tianjin: Another look.' *Journal of Chinese Linguistics* 15.2 (1987): 247–273