An Optimality Theory Perspective on Mandarin Tone 3 Sandhi

Dr. Hui Yin
Department of English, Xi’an Jiaotong-Liverpool University, China.

*Corresponding Author: Dr. Hui Yin, Department of English, Xi’an Jiaotong-Liverpool University, China.

ABSTRACT
This paper aims to provide an Optimality Theory (OT) account of Mandarin tone 3 sandhi. Some linguists have noticed that there are obvious phonetic differences between Mandarin sandhied tone 3 and tone 2 and that they are not neutralized. Traditional analyses viewing tone 3 sandhi as categorical change fail to account for differences between sandhied tone 3 and tone 2. In this paper a constraints-based analysis within the Optimality Theory (OT) framework is developed to look into Mandarin tone 3 sandhi and to explain differences between sandhied tone 3 and tone 2. A new analysis claiming that tone 3 sandhi is a tonal reduction is proposed and justified within the OT framework. The OT device of constraint conflicts permits the present analysis of tone 3 sandhi as a resolution of conflicts between a general principle of markedness and a general principle of tonal faithfulness. This approach provides straightforward descriptions and motivations for Mandarin tone 3 sandhi.

Keywords: tone, sandhi, Optimality Theory, constraint, Mandarin

INTRODUCTION
There are four basic tone categories in Mandarin Chinese. Every stressed syllable belongs to one of the four, although their actual realizations may vary according to context (Ch’en et al., 1994). When the four tones are applied to the same syllable, four different lexical items may result (Wang & Norval, 1997). The following is the classic example of tonal contrasts:

(1) syllable      tone    gloss
    ma      tone 1: high level  ‘mother’
    ma      tone 2: high rise   ‘hemp’
    ma      tone 3: fall-rise    ‘horse’
    ma      tone 4: falling     ‘scold’

In analyses of the Mandarin tonal system, the voice range is divided into five levels: upper, upper middle, middle, lower middle and lower. The numbers from 1 to 5 can be used to designate these levels, where 1 represents the lowest comfortable pitch of the vocal range and 5 represents the highest. The first tone (55) is high and level. It is pitched near the top of the comfortable voice range. The second tone (35) starts around the middle of the voice range 3 and rises straight towards the level of the first tone 5.

The third tone (214) begins near the bottom of the comfortable voice 2, proceeds to the bottom 1, and then upward to end above the middle 4. The fourth tone (51) begins at the top of the comfortable range 5 and falls quickly to the bottom 1 (Ch’en et al., 1994).

A well-known phenomenon in Mandarin phonology is tonal variations according to different phrasal contexts, among which tone 3 sandhi is the most noticeable and most complicated. In my analysis, tone 3 sandhi can be regarded as a kind of tonal reduction. I will show that a set of OT constraints can be employed to accommodate Mandarin tone 3 sandhi and that the analysis proposed in this paper is able to account for the difference between the sandhied tone 3 and lexical tone 2. In the next section, I will briefly discuss the phenomenon of Mandarin tone 3 sandhi.

TONE 3 SANDHI

In rule-based analyses, the basic rule of Mandarin tone 3 sandhi can be simply stated as in (2):
An Optimality Theory Perspective on Mandarin Tone 3 Sandhi

(2) 3---->[2]/___3

When one third tone syllable is immediately followed by another, the first of the two changes from a “falling-rising tone [214] into high rising [35]” (Chen, 2000:20) i.e. tone 2, which is indicated by brackets in rule-based analyses throughout the paper. For example, in (3) the second syllable jiu ‘wine’ changes from tone 3 to tone 2 according to traditional analyses.

(3) mai3 jiu3 ---> mai[2] jiu3

buy wine

In Mandarin, the domain for the third sandhi rule application can be larger than two syllables as the example in (4) shows.


buy good wine

In (4), there are three tone 3 syllables in a string, and all the third tones except for the last one change to the second tones by the third tone sandhi rule according to the convention assumed by Chomsky and Halle (1968).

However, some linguists (e.g. Hockett, 1950; Martin, 1963) claim that the changed third tone does not really become a tone 2 but rather a new tone (tone 5). Phonetically, there are some differences between the changed tone 3 and the underlying tone 2 though it sounds like an underlying tone 2 (e.g. Wang & Li, 1967). There are a number of linguists who have demonstrated this kind of phonetic differences (their data mainly come from Beijing Mandarin or mainland Mandarin). Kratochvil (1984) claims that tone 3 + tone 3 seems to be quite distinct from tone 2 + tone 3 phonetically and his experiment (1984) shows that the first tone 3 in a two tone 3 sequence is more often classified as tone 3 rather than tone 2. Zee (1980) reports that the mean fundamental frequency (F0) for sandhied tone 3 is 17.5 Hz lower than that of tone 2 and he concludes that sandhied tone 3 is not identical to underlying tone 2. Xu (1997) also has data which clearly indicate that the third tone sandhi results in a tone lower than tone 2. Yuan and Chen report (2014:218) that the sandhied tone was “different from the lexical rising tone (Tone 2) in disyllabic words in two measures: the magnitude of the F0 rise and the time span of the F0 rise”. However, these phonetic differences have received little attention in traditional third tone sandhi analyses.

Moreover, traditional rule-based approaches fail to answer some fundamental questions regarding tone 3 sandhi such as the following:

- How can we account for the differences between the so-called derived tone 2 from tone 3 sandhi and the lexical tone 2?
- Why does it happen that the first tone 3 undergoes tone 3 sandhi instead of the last one in a 33 tone sequence?

In Mandarin phonology literature (e.g. Zhang, 1997; Duanmu, 1999 & 2000; Lin, 2002; Yip, 2002; Yin 2003; Wee, 2008 & 2010; Lee-Schoenfeld & Kandybowicz 2009), there are some OT analyses regarding tone 3 sandhi, and these analyses mainly deal with tone 3 sandhi domains and/or the role of speech rate in tone 3 sandhi. However, the two afore-mentioned basic questions have yet to be answered. Moreover, studies on Mandarin third tone sandhi have focused almost exclusively on the context of a string on two or more consecutive third tones. It has been noticed that when a third tone follows a tone other than a third tone, it undergoes sandhi as well. In that case, a third tone will retain its first half (that is [21] if numbers are used to represent it) as (5) illustrates.

(5) mai3 shu1-> mai [first half of tone 3 or [21] shu]

buy book

In (5), tone 3 follows tone 1 and it retains its first half rather than its second half. This kind of third tone modification in the context of a third tone 3 plus a non-third tone has been ignored by traditional analyses. Then why does the underlying third tone retain its falling portion (the first half) instead of its rising portion (the second half) when followed by a tone other than a third tone? Like the two afore-mentioned questions, this question remains unanswered as well. This paper aims to answer these three questions and offer a unified account of the two third tone sandhi phenomena from an Optimality Theory perspective.

THE PRESENT ANALYSIS AND CONSTRAINTS

In this section, I will propose an OT analysis to address the third tone sandhi in Mandarin Chinese. With markedness and faithfulness constraints within the OT framework (Prince & Smolensky, 2004; McCarthy & Prince, 1993), this analysis provides a straightforward account for the third tone sandhi. A set of basic constraints related to the third tone sandhi will be addressed as well.
The present analysis is based on the OT framework. An argument in favor of OT could be constructed. The OT device of constraint conflict permits an analysis of tone 3 sandhi as a resolution of conflict between a general principle of effort minimization and a general principle of tonal faithfulness. This could largely explain the loss of the falling portion of tone 3 in a 33 sequence and the loss of its rising portion in other contexts. As mentioned previously, the full third tone contour is 214 if numbers are used to represent it. Thus, I will use the notations of [214] for a full tone 3, [14] for tone 3 with loss of the falling portion and [21] for tone 3 with loss of the rising portion in the OT analysis in this paper. Therefore, the present analysis treats the third tone sandhi as a tonal reduction phenomenon rather than categorical tonal change. Naturally the sandhiied tone 3 which is a reduction of a full tone 3 is different from underlying tone 2 phonetically.

The basic idea of OT is that “each linguistic output is optimal, in the sense that it incurs the least serious violations of a set of conflicting constraints” (Kager, 1999:8). Outputs are results of conflicts between markedness and faithfulness. Surface variations could result from rankings of relevant constraints. The OT framework is based on markedness constraints and faithfulness constraints. Such constraints are violable, but violation must be minimal. “Markedness constraints require that output forms meet some criterion of structural well-formedness” (Kager, 1999: 9). With regard to third tone sandhi in Mandarin, one of the important markedness constraints is final lengthening/non-final shortening in a phonological phrase as in (6).

(6) Final-L/N-Final-S

In a phonological phrase, the final syllable should be lengthened and the non-final syllable(s) should be shortened.

This markedness constraint is well motivated cross-linguistically (e.g., Nooteboom, 1997, Duanmu, 2000). In the case of Mandarin Chinese, the final syllable in a phonological phrase usually indicates a major boundary.

Another markedness constraint active in the third tone sandhi is the effort minimization constraint:

(7) *ComCT/ShortS

Complex contour tones should be prohibited in shortened syllables.

In Mandarin, tone 3 consists of two parts: a falling portion and a rising part. Thus, it can be regarded as a complex contour tone. The tone 3 sandhi process could be characterized as a tonal reduction (loss of the first part—the falling portion of tone 3 in a 33 sequence) in contexts where the syllable has a shortened or non-lengthened duration. In other contexts, tone 3 will lose its rising portion (second part) due to a markedness constraint: the T-POLARITY constraint, which requires an initial tone to be followed by an opposite tone (Duanmu 1999). The phenomenon for a tone 3 to lose its second half will be dealt with in detail later. Duanmu (1999) claims that the duration of the third tone is longer than that of other tones. If enough duration cannot be maintained, the full third tone as a complex contour tone would be hard to be realized. This is a consequence of the effort minimization principle—the shortened duration increases the effort cost of achieving all 3 tonal targets. So it is reasonable to assume that the tone 3 sandhi is a kind of tonal reduction process.

Faithfulness constraints require that “outputs preserve the properties of their basic (lexical) forms, requiring some kind of similarity between the output and the input” (Kager, 1999). The relevant faithfulness constraint in third tone sandhi is one of the correspondence constraints.

(8) MAX (contour)

Correspondent segments in input and output have identical values for tone contours. In the case of the third tone, the contour refers to both the falling part (I use contour numbers 21 to represent it) and the rising part (I use contour numbers 14 for this part).

Thus we can derive two sub-constraints of correspondence regarding tone 3:

(a) MAX (21)
Input segments must have output correspondents in terms of the falling part of tone 3.

(b) MAX (14)
Input segments must have output correspondents in terms of the rising part of tone 3

In this paper, the numbers [14] are used to represent the underlying representation of the rising portion of tone 3 but besides [14] its actual surface realization could be [24], [34] or any contour between 1 and 4 as allophones due to some factors such as further reductions in non-final contexts or individual speakers’
An Optimality Theory Perspective on Mandarin Tone 3 Sandhi

differences. In any way, these allophones maintain the basic characteristics of its rising portion of the third tone. For the sake of simplicity, I use [14] to characterize the rising portion of the third tone in this paper.

In this section, a set of constraints has been proposed and these constraints play a role in Mandarin tone sandhi. The next section will address how these constraints are applied to Mandarin tone 3 sandhi.

APPLICATION OF THE CONSTRAINTS

In Mandarin, when Final-L/Non-Final-S and *ComCT/ShortS rank higher than Max (contour), third tone sandhi rule will apply as the following example hao3 jiu3 ‘good wine’ illustrates. (In the input, I will specify which tone it is and in the output I will use contour numbers to represent it, e.g. [214] for a full tone 3). Following traditions of OT, relevant constraints are listed on the top row horizontally in a descending ranking from left to right. However, the dashed line shows that there is no crucial ranking between the constraint on the right and the one on the left. The input is shown in the top left cell. The output candidates are indicated in the first column. An asterisk mark ‘*’ indicates a violation of a constraint heading the column and an exclamation mark ‘!’ indicates a violation that rules out a candidate. The optimal candidate is marked by the index ☞.

The optimal candidate is marked by the index ☞. Shaded cells mean that they are no longer relevant for the evaluation.

Table 1. Evaluation of candidates for /hao3 jiu3/

<table>
<thead>
<tr>
<th>Input /hao3 jiu3/</th>
<th>Final-L/N-Final-S</th>
<th>*ComCT/ShortS</th>
<th>Max (21)</th>
<th>Max (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>☞ a. (hao[14] jiu[214])</td>
<td>☞</td>
<td>☞</td>
<td>☞</td>
<td>☞</td>
</tr>
<tr>
<td>b. (hao[14] jiu[14])</td>
<td>*!</td>
<td>☞</td>
<td>☞</td>
<td>☞</td>
</tr>
<tr>
<td>c. (hao[214] jiu[214])</td>
<td>☞</td>
<td>*!</td>
<td>☞</td>
<td>☞</td>
</tr>
<tr>
<td>d. (hao[214] jiu[214])</td>
<td>☞</td>
<td>☞</td>
<td>☞</td>
<td>☞</td>
</tr>
</tbody>
</table>

In this section, a set of constraints has been proposed and these constraints play a role in Mandarin tone sandhi. The next section will address how these constraints are applied to Mandarin tone 3 sandhi. For example, in Mandarin usually a nominal compound of more than two morphemes constitutes a phonological phrase as in (9).

‘exhibition hall’

In (9), the lexical tones for the tri-syllabic nominal compound are all third tones. In its surface realization, except for the final third tone, the non-final third tones change to half third tones which are represented by numbers [14] here. Table 2 shows the rankings of relevant constraints and evaluations of competing candidates.

Table 2. Evaluation of candidates for /zhan3 lan3 guan3/

<table>
<thead>
<tr>
<th>Input /zhan3 lan3 guan3/</th>
<th>Final-L/N-Final-S</th>
<th>*ComCT/ShortS</th>
<th>Max (21)</th>
<th>Max (14)</th>
</tr>
</thead>
</table>

As Table 2 indicates that candidate (b) zhan[14] lan[14] guan[14] violates the higher ranked constraint Final-L/N-Final-S and it cannot win the competition. In candidate (c) zhan[14] lan[214] guan[214], the second syllable is a non-final syllable and it gets shortened. This shortened syllable is within the domain of a phonological phrase but it keeps its full tone.
An Optimality Theory Perspective on Mandarin Tone 3 Sandhi

which is a complex contour tone. Thus, it violates the higher ranked constraint *ComCT/ShortS and cannot survive either. Candidate (a) zhan[14] lan[14] guan[214] only violates the lower ranked constraint MAX (21) and is the winning candidate.

Phonological phrases are not only sensitive to morphological structure but also to syntactic structure. For example, a verb and its object form a phonological phrase in the application of the third tone sandhi rule as in (10).

(10) mai3 hao3 ma3 ---> mai[14] hao[14] ma3

Table 3. Evaluation of candidates for /mai3 hao3 ma3/

<table>
<thead>
<tr>
<th>Input /mai3 hao3 ma3/</th>
<th>Final-L/N-Final-S</th>
<th>*ComCT/ShortS</th>
<th>MAX (21)</th>
<th>MAX (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>c. (mai[214] hao[14] ma[214])</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Table 3, the final syllable ma ‘horse’ in candidate (b) mai[14] hao[14] ma[14] is not lengthened and it does not abide by the Final-L/N-Final-S constraint which is ranked higher. Thus, this candidate cannot be a winning candidate. In candidate (c) mai[214] hao[14] ma[214], the non-final syllable mai ‘buy’ keeps its full third tone which is a complex contour tone. Therefore, it violates the higher ranked constraint *ComCT/ShortS and cannot be an ideal candidate. Candidate (a) mai[14] hao[14] ma[214] only violates the lower ranked constraint MAX (21) and compared with other candidates it occurs the least violation. As a result, this candidate survives as an ideal candidate.

In Mandarin, subject and predicate constitute a major syntactic boundary. In normal speech, they do not form a phonological phrase, and thus, normally the third tone sandhi rule does not apply between subject and predicate as (11) shows.

(11) xiao3 ma3 hao3->(xiao[14] ma[214]) (hao[214]) small horse

’small horses are good’

Examples (9), (10) and (11) all have three syllables; however, examples (9) and (10) constitute only one phonological phrase each while example (11) forms two phonological phrases indicated by parentheses. Table 4 shows the rankings of relevant constraints and evaluations of competing candidates.

Table 4. Evaluation of candidates for /xiao3 ma3 hao3/  

<table>
<thead>
<tr>
<th>Input / xiao3 ma3 hao3/</th>
<th>Final-L/N-Final-S</th>
<th>*ComCT/ShortS</th>
<th>MAX (21)</th>
<th>MAX (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (xiao[14] ma[214]) (hao[214])</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (xiao[14] ma[14]) (hao[214])</td>
<td>*!</td>
<td></td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>c. (xiao[214] ma[214]) (hao[214])</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Therefore, it violates the higher ranked constraint *ComCT/ShortS and cannot survive as an ideal candidate. In candidate (a) xiao[14] ma[214] hao[214], the non-final syllable xiao ‘small’ undergoes the sandhi by losing its first half tone. This candidate only violates the lower ranked constraint MAX (21) and becomes the winning candidate eventually.

Then how can we account for the fact that the rising part rather than the falling part drops when tone 3 precedes any other tone? Here still
An Optimality Theory Perspective on Mandarin Tone 3 Sandhi

another markedness constraint is at work. I adopt Duanmu’s constraint: T-POLARITY (1999) to explain this phenomenon (and also explain the loss of the falling part rather than the rising part of tone 3 in 33 sequences):

(12) T-Pol

A tonal polarity constraint requires an initial tone to be followed by an opposite tone, that is, a low tone should be followed by a high tone and vice versa.

In Mandarin tonal system, tone 3 is a complex tone which could be viewed as a low tone (falling part) plus a high tone (rising part) whereas all the other three tones basically can be regarded as high tones. Duanmu (1999:23) points out that a monosyllabic tone 3 in Mandarin usually surfaces as two parts, which ends in a H (high tone). Chao (1933:132) also claims that a monosyllabic tone 3 “often breaks into two parts” with a glottal stop in between. As for the other tones, tone 1 (55) and tone 2 (35) as a whole can be viewed as high tones. In tone 4 (51), the first portion beginning with (5) is a high tone. So when tone 4 follows tone 3, the most adjacent portion to tone 3 is a high tone in a tone 3 and tone 4 sequence. Thus, in the sequence of tone 3 plus a tone other than tone 3, the non-final tone 3 needs to keep its falling part which is a low tone in order to satisfy the T-POLARITY constraint.

The following table illustrates that tone 3 retains its falling part rather than its rising part in the example of hao3 shu1 ‘good book’:

Table 5. Evaluation of candidates for /hao3 shu1/

<table>
<thead>
<tr>
<th>Input /hao3 shu1/</th>
<th>Final-L/N-Final-S</th>
<th>*ComCT/ShortS</th>
<th>T-Pol</th>
<th>MAX (21)</th>
<th>MAX (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a (hao[21] shu[55])</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b (hao[14] shu[55])</td>
<td></td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Table 5, candidate (b) violates the constraint T-POLARITY while (a) abides by the T-POLARITY constraint. It can be seen from the table that candidates (a) incurs the least serious violations though it violates the lower ranked MAX(14) constraint and accordingly, it is the ideal candidate. Then how can we account for the phenomenon that in a two tone 3 sequence the first tone 3 should retain its rising portion instead of its falling portion? With the same ranking as in Table 5, we can derive the winning output straightforwardly.

Table 6. Evaluation of candidates for /hao3 ma3/ ‘good horse’

<table>
<thead>
<tr>
<th>Input /hao3 ma3/ ‘good horse’</th>
<th>Final-L/N-Final-S</th>
<th>*ComCT/ShortS</th>
<th>T-Pol</th>
<th>MAX (21)</th>
<th>MAX (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a (hao[14] ma[21])</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b (hao[21] ma[21])</td>
<td></td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As illustrated in Table 6, in the example of hao3 ma3 ‘good horse’, the second syllable ma3 ‘horse’ is realized as a full tone 3 to begin with a low tone (21) since it gets lengthened in the final position and thus, it should be preceded by an opposite tone, a high tone to satisfy the T-POLARITY constraint. As a result, the first tone 3 in a 33 sequence should keep its rising portion rather than its falling portion to be a winning candidate.

CONCLUSIONS

This study investigates the most complicated tonal change in Mandarin: the third tone sandhi. A constraints-based analysis within the OT framework is developed in this paper to look into third tone sandhi in Mandarin. With proper rankings of markedness constraints and faithfulness constraints, this approach provides straightforward descriptions for third tone sandhi and can cover all the data including those which traditional approaches fail to account for. Thus this paper has demonstrated that with the constraints-based analysis within the OT framework, third tone sandhi in Mandarin can be fully accounted for.

In the present analysis, tone 3 sandhi sandhi is viewed as tonal reductions instead of categorical changes. Since the application of the tone sandhi leads to a half tone it would be reasonable to analyze the sandhielined tone as phonetic rather than a phonological variant of the unsandhielined form. Tone 3 which has undergone the sandhi retains its rising portion (second half) rather than neutralizes to tone 2, and thus, it is naturally different from underlying tone 2 phonetically.

In a word, according to OT, constraints are violable and the optimal candidate should incur
An Optimality Theory Perspective on Mandarin Tone 3 Sandhi

least serious violations of these constraints. In fact, outputs are results of conflicts between markedness constraints and faithfulness constraints. Surface variations could result from proper rankings of relevant constraints.

REFERENCES

AUTHOR’S BIOGRAPHY
Dr. Hui Yin is Associate Professor and Director of English and Communication Program at the Department of English, Xi’an Jiaotong-Liverpool University. He is also in the Board of Directors of Linguistic Association of Canada and the United States. Before joining Xi’an Jiaotong-Liverpool University, Dr. Yin was Professor at Xiamen University of Technology, China; Linguist at Zi Corporation, Canada; Associate Professor and Director at the Institute of College English, Zhejiang University, China; Visiting Scholar, Central College, USA. His main research areas lie in Cognitive Linguistics, Corpus Linguistics, Phonology, Syntax, Language Teaching and Learning.