Proceedings of the 23rd North American Conference on Chinese Lingusitics (NACCL-23), 2011.
Volume 1, edited by Zhuo Jing-Schmidt, University of Oregon, Eugene. Pages 138-155.

# Variation in Tone 3 Sandhi: The case of prepositions and pronouns*

Chiung-Yao Wang and Yen-Hwei Lin
*Michigan State University*

This paper examines variation in Tone 3 Sandhi (henceforth T3S) with a focus on the case of prepositions and pronouns. Prepositions have been observed to produce an additional T3S surface pattern. Pronouns are found to behave differently from nouns in T3S application. The major approach for analyzing these cases is primarily the pre-OT (Optimality Theory) derivational process through which multiple T3S patterns surface in the output. Zhang (1997) analyzes examples with prepositions and pronouns within the OT framework. However, her approach requires a two-step two-tableau process. We adopt Coetzee's (2006) OT variation model to account for (i) the behavior of prepositions and pronouns in T3S and (ii) the variability in T3S surface forms, two of the challenging aspects in analyzing T3S application in sentences. The advantages of our re-analysis are twofold: multiple surface patterns are obtained in one step, and the relative frequency of surface patterns is indicated.

## 1. Introduction

Mandarin Tone 3 Sandhi (henceforth T3S) has been the most extensively studied tone sandhi phenomenon in Mandarin Chinese. T3S changes a Tone 3 (T3) to a Tone 2 (T2) before another T3. This simplified description of T3S is given in (1).

(1) Mandarin T3S:
T3 → T2/__ T3
(T2 = mid rising tone (35)
T3 = dipping tone (214) phrase finally and low tone (21) elsewhere)

The rule describe in (1) is deceptively simple as how T3S applies becomes rather complicated in cases where there are more than two T3s in a word or phrase (Lin 2007:204). In phrases and sentences where there are multiple adjacent T3's, T3S often exhibits variation by allowing more than one surface pattern depending on syntactic structure and how it is prosodified as shown in (2).

* We thank the participants of NACCL-23 for their comments and questions.

| | | | | | |
|---|---|---|---|---|---|
| (2) | [*Lao* | *Li*] | [*mai* | [*hao* | *jiu*]] |
| | Lao | Li | buy | good | wine |
| | 'Old Li buys good wine.' | | | | |
| | T3 | T3 | T3 | T3 | T3 |
| → | (**T2** | T3) | (T3) | (**T2** | T3) |
| → | (**T2** | **T2** | T3) | (**T2** | T3) |
| → | (**T2** | **T2** | **T2** | **T2** | T3) |

In accounting for variation in T3S, the primary approach has been the pre-OT derivational analysis (Cheng 1973, Shih1986, 1997, Chen 2000, Duanmu 2000/2007, Lin 2007). The additional patterns are often considered to be present in faster or casual speech (e.g. Cheng 1973, Shih1986, 1997, Zhang 1997, Chen 2000, Lin 2007), but are considered to be just possible variants unrelated to speech rates in some studies (e.g. Duanmu 2000/2007). In either case, the question is how variation in Mandarin T3S can be modeled in Optimality Theory (OT).

In this paper, we analyze simple examples containing a prepositional phrase (PP) with or without a pronoun by adopting Coetzee's (2006) OT variation model. The next section (§2) presents the data of T3S and previous analyses, and §3 provides background for modeling variation in OT. Our proposed OT analysis is presented in §4. The concluding section (§5) offers remarks on theoretical implications and suggests what can be done in future T3S studies.

## 2. T3S: the case of prepositions and pronouns

In this section, we first look at the T3S data, focusing on examples containing prepositions and pronouns. Previous analyses are presented, both a pre-OT analysis (Shih 1986, 1997, Chen 2000, Lin 2007) and an OT analysis (Zhang 1997).

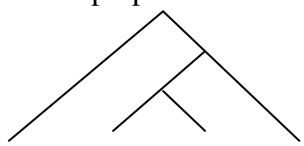
### 2.1. The data

Consider the examples in (3) first. ST3[1] is grammatical in (3b) but it is not in (3a) although the sentences have the same branching structure. The fact that ST3 surfaces in (3b) is not because the lexicon *bi* 'than' is a special case. The same pattern is found in other prepositions (e.g. *wang* 'toward' in [*Gou* [[*wang bei*] *zou*]] 'The dog walks toward the north'). The fact that prepositions tend to be prosodically weak is responsible for the additional pattern.

[1] ST3 is ungrammatical for some audience at NACCL-23. One participant pointed out that the judgment of grammaticality may result from regional differences (which we agree), and the absence of the syllable "*gou*" in T2 (sandhi tone for *gou* 'dog') may play a role in speakers' judgment. However, 苟且偷生 *gou qie tou sheng* 'to drift and live without purpose' T3T3T1T1→ **T2**T3T1T1 is found, so the lack of an independent word for the syllable *gou* with T2 does not prohibit a sandhi tone to appear on the surface.

(3) Same branching structure with or without a preposition (cf. Zhang 1997)

a. Without a preposition

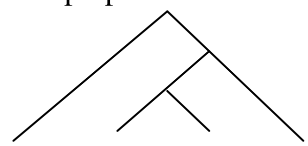


| | | | | |
|---|---|---|---|---|
| Ma | hen | shao | hou | |
| horse | very | rarely | roar | |

'A horse rarely roars.'

| | | | | |
|---|---|---|---|---|
| 3 | 3 | 3 | 3 | UT |
| (3) | (2 | 2 | 3) | ST1 |
| (2 | 2 | 2 | 3) | ST2 |
| *(2 | 3) | (2 | 3) | |

b. With a preposition



| | | | | |
|---|---|---|---|---|
| Gou | **bi** | ma | xiao | |
| dog | than | horse | small | |

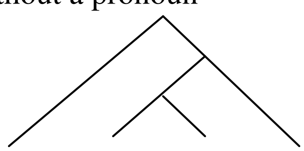
'A dog is smaller than a horse.'

| | | | | |
|---|---|---|---|---|
| 3 | 3 | 3 | 3 | UT |
| (3) | (2 | 2 | 3) | ST1 |
| (2 | 2 | 2 | 3) | ST2 |
| **(2** | **3)** | **(2** | **3)** | **ST3** |

(UT= Underlying tones; ST1= surface tones; Surface pattern 1; ST2= Surface pattern 2...)

(4) Same branching structure containing a PP with or without a pronoun

a. Without a pronoun

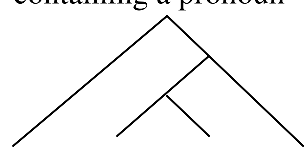


| | | | | |
|---|---|---|---|---|
| Gou | **bi** | ma | xiao | |
| dog | than | horse | small | |

'A dog is smaller than a horse.'

| | | | | |
|---|---|---|---|---|
| 3 | 3 | 3 | 3 | UT |
| (3) | (2 | 2 | 3) | ST1 |
| (2 | 2 | 2 | 3) | ST2 |
| **(2** | **3)** | **(2** | **3)** | **ST3** |

b. PP containing a pronoun



| | | | | |
|---|---|---|---|---|
| Gou | **bi** | **wo** | xiao | |
| dog | than | I | small | |

'A dog is smaller than me.'

| | | | | |
|---|---|---|---|---|
| 3 | 3 | 3 | 3 | UT |
| (3) | (2 | 2 | 3) | ST1 |
| (2 | 2 | 2 | 3) | ST2 |
| *(2 | 3) | (2 | 3) | |

Like prepositions, pronouns are prosodically weak and their behavior differs from nouns. The sentences in (4a) and (4b) have the same branching structure containing a PP; the only difference lies in what follows the preposition: a noun or a pronoun. ST3 is grammatical in (4a) where a noun follows the preposition, but the same surface pattern is banned in (4b) where a pronoun follows the preposition.

The data in (3) and (4) show two challenges in analyzing T3S. First, any T3S analysis has to account for the variation as we see that one input is mapped onto more than one output. Second, sentences of the same syntactic branching structure do not always exhibit the same T3S surface patterns. That is, syntax alone is insufficient in accounting for the differences.

## 2.2. Pre-OT analysis

The contrasting patterns in sentences in (3) and (4) have been accounted for through different syntax-based prosodic parses. Surface variants are obtained by derivational steps (e.g. Shih 1986, 1997, Chen 2000: Ch.9, Duanmu 2000/2007: Ch.11, Lin 2007: Ch.9), as shown in (5) – (7). The derivational process follows that in Lin (2007).

(5) Without a preposition

| | | | | |
|---|---|---|---|---|
| [*Gou* | [[*hen* | *shao*] | hou.]] | |
| dog | very | rarely | roar | 'A dog rarely roars.' |
| T3 | T3 | T3 | T3 | UT |

Normal speech variant

| | | | | |
|---|---|---|---|---|
| T3 | T3 | T3 | T3 | Word level: Not applicable |
| T3 | (**T2** | T3) | T3 | Phrase level: Disyllabic foot, T3S |
| T3 | (T2 | **T2** | T3) | Phrase level: Incorporation, T3S |
| (T3 | T2 | T2 | T3) | Phrase level: Incorporation, no T3S; ST1 |

Fast speech or alternative variant

| | | | | |
|---|---|---|---|---|
| (**T2** | **T2** | **T2** | T3) | One prosodic domain, T3S; ST2 |

In (5), at the Word level, T3S is not applicable. At the Phrase level, a disyllabic foot is parsed for the smallest domain *hen shao* 'very seldom', and T3S applies. When the verb *hou* 'roar' is incorporated, T3S applies again. Finally, when *ma* 'horse' is incorporated, T3S does not apply because there are no adjacent T3's. The first surface pattern derived is (T3**T2T2**T3). ST2 is a pattern in fast speech where a larger domain is parsed and T3S applies from left to right (Shih 1986, 1997, Chen 2000, Lin 2007), or it is an alternative pattern unrelated to fast speech (Duanmu 2000/2007). In (6), the branching structure is the same as that in (5), but in (6) there is a preposition which can, but does not have to cliticize.

(6) With a preposition, PP containing a noun

| | | | | |
|---|---|---|---|---|
| [*Gou* | [[*bi* | *ma*] | *xiao*.]] | |
| dog | than | horse | small | 'Dogs are smaller than horses.' |
| T3 | T3 | T3 | T3 | UT |

No cliticization

| | | | | |
|---|---|---|---|---|
| T3 | T3 | T3 | T3 | Word level: Not applicable |
| T3 | (**T2** | T3) | T3 | Phrase level: Disyllabic foot, T3S |
| T3 | (T2 | **T2** | T3) | Phrase level: Incorporation, T3S |
| (T3 | T2 | T2 | T3) | Phrase level: Incorporation, no T3S; ST1 |

Fast Speech or alternative variant

| | | | | |
|---|---|---|---|---|
| (**T2** | **T2** | **T2** | T3) | One prosodic domain, T3S; ST2 |

Cliticization

| | | | | |
|---|---|---|---|---|
| T3 | T3 | T3 | T3 | Word level: Not applicable |
| (**T2** | T3) | T3 | T3 | Cliticization, T3S |
| (T2 | T3) | (**T2** | T3) | Phrase level: disyllabic foot, T3S; ST3 |

In (6), when the preposition does not cliticize, the derivation is the same as that in normal speech variant in (5). In the case of ST1 in (6), the preposition is not treated differently from other word categories. ST2 is the fast speech or alternative variant when a larger domain is parsed and T3S applies from left to right. When the preposition cliticizes, ST3 surfaces. At the Word level, T3S is not applicable. Next, the preposition cliticizes to the host on its left, and T3S applies. Finally, a disyllabic foot is formed for the rest of the syllables, and T3S applies again. ST3 (**T2**T3)(**T2**T3) is thus derived. This same surface pattern is ungrammatical in (7) where a pronoun follows the preposition.

(7) With a preposition, PP containing a pronoun

| | | | | |
|---|---|---|---|---|
| [*Gou* | [[*bi* | *wo*] | *xiao*]] | |
| dog | than | I | small | 'Dogs are smaller than me.' |
| T3 | T3 | T3 | T3 | UT |

No cliticization

| | | | | |
|---|---|---|---|---|
| T3 | T3 | T3 | T3 | Word level: Not applicable |
| T3 | (**T2** | T3) | T3 | Phrase level: Disyllabic foot, T3S |
| T3 | (T2 | **T2** | T3) | Phrase level: Incorporation, T3S |
| (T3 | T2 | T2 | T3) | Phrase level: Incorporation, no T3S; ST1 |

Fast Speech or alternative variant

| | | | | |
|---|---|---|---|---|
| (**T2** | **T2** | **T2** | T3) | One prosodic domain in fast speech, T3S; ST2 |

Cliticization

| | | | | |
|---|---|---|---|---|
| T3 | T3 | T3 | T3 | Word level: Not applicable |
| T3 | (**T2** | T3) | T3 | Phrase level: Cliticization, T3S |
| T2 | (T2 | **T2** | T3) | Phrase level: Incorporation, T3S |
| (T3 | T2 | T2 | T3) | Phrase level: Incorporation, no T3S; ST3 (= ST1) |

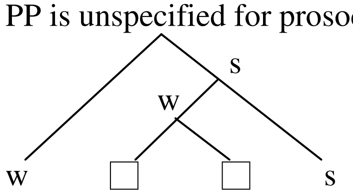
In (7), the derivational steps for ST1 and ST2 are the same as those in (6). For ST3, at the Word level, T3S is not applicable. At the Phrase level, the pronoun *wo* 'I' cliticizes to the preposition *bi* 'than', and T3S applies. According to Shih (1997:110), a pronoun is weaker than a preposition and will cliticize leftwards onto the preposition, and as a two-syllable structure, it will not cliticize further because it is heavy enough. When the last syllable *xiao* 'small' is incorporated in the final step, T3S does not apply as there are no adjacent T3's. The derived pattern ST3 is T3**T2T2**T3, identical to the pattern in ST1, even though the derivational processes of ST1 and ST3 differ.

As shown in (5) – (7), the pre-OT analysis accounts for all the surface patterns in the contrasting data derivationally. We now turn to Zhang (1997) who offered an OT account for the data.

### 2.3. Zhang's (1997) two-step two-tableau OT analysis

A major difference in Zhang's analysis (1997) from the previous studies is her use of unspecified *strong/weak* in the metrical structure for preposition phrases in accounting for multiple T3S outputs. Zhang adopts Cinque's (1993) Null Theory of Phrase Stress and she takes strong/weak stress in Cinque's theory as the strong/weak constituent strength (Zhang 1997:304). According to Zhang, a prepositional phrase is unspecified for its strength, and it can be either strong or weak (Zhang 1997:304-305). The sentence in (8) shows unspecified strength for the PP.

(8) PP is unspecified for prosodic strength, so it can be either strong or weak.




| | | | | | |
|---|---|---|---|---|---|
| | **s** | **w** | | → | possibility 1 |
| | **w** | **s** | | → | possibility 2 |
| *Gou* | *bi* | *ma* | *xiao*. | | |
| dog | than | horse | small | 'Dogs are smaller than horses.' | |
| T3 | T3 | T3 | T3 | UT | |

A sentence that has a PP, unspecified for its prosodic strength, then enters the two inputs in two separate OT tableaux for evaluation. It should be noted that the optimal outputs for PP-strong-weak and PP-weak-strong may completely overlap, partially overlap, or do not overlap. Six constraints used in Zhang's (1997) analysis are given in (9).

(9) Constraints used in Zhang (1997:306-308)

a. *33 (no sequential third tones): No adjacent third tones are allowed.

b. Cl (Clitic Dependency): A clitic cannot be separated from the TS domain of the preceding verb or preposition head.

c. Max (Maximal Domain): The maximal TS domain is two syllables in normal speaking rate, but larger in more casual or faster style.
d. Align-Di-L (Disyllabic Constituent Alignment): Align the left side of a TS domain with the left side of a disyllabic constituent when two or more TS domains occur.
e. PTAS (Parse UT of an Absolutely Strong Node): The underlying tone of a strong constituent which is not dominated by any *w* node must be parsed.
f. PTRS (Parse UT of a Relatively Strong Node): The underlying tone of a strong constituent which is dominated by at least one *w* node must be parsed.

*Max* maintains that the ideal length of a prosodic domain is two syllables at the normal speech rate (Zhang 1997:308). PTAS and PTRS ensure a strong element to be faithful to its underlying tone (Zhang 1997:306). The constraint Align-Di-L, according to Zhang's definition, is irrelevant when there is only one T3S domain. Zhang uses the symbol □ to denote the unspecified constituent strength, which could be either *strong* or *weak*. We now turn to two examples of her analysis in (10) – (11).

(10) Zhang's two-input-two-tableau analysis: PP containing a noun
[*Gou* [[*bi* *ma*] *xiao*]]
dog than horse small ‘Dogs are smaller than horses.’

a. PP is ws (Zhang 1997:314)

| wwsS 3333 | PTAS | *33 | PTRS | Align-Di-L | Max |
|---|---|---|---|---|---|
| a. (223)(3) | | *! | | * | * |
| b. (23)(22) | *! | | * | * | |
| ☞c. (3)(223) | | | * | | * |
| ☞d. (2223) | | | * | | * |
| e. (23)(23) | | | * | *! | |

b. PP is sw (Zhang 1997:315)

| wswS 3333 | PTAS | *33 | PTRS | Align-Di-L | Max |
|---|---|---|---|---|---|
| a. (223)(3) | | *! | * | * | * |
| b. (3)(223) | | | * | | *! |
| c. (2223) | | | * | | *! |
| ☞d. (23)(23) | | | | * | |

In (10a) where PP is *ws*, Candidate (c) (T3)(**T2T2**T3) and Candidate (d) (**T2T2T2**T3) are optimal outputs. In (10b) where PP is sw, Candidate (d) (**T2**T3)(**T2**T3) is the optimal output. The optimal outputs from both tableaux are final winners. All three surface patterns for *Gou bi ma xiao* ‘Dogs are smaller than horses’ are accounted for.

(11) Zhang's two-input-two-tableau analysis: PP containing a pronoun
[*Gou* [[*bi* *wo*] *xiao*]]
dog than I small 'Dogs are smaller than me.'

a. PP is *ws* (Zhang 1997:316)

| wwsS 3333 | PTAS | *33 | Cl | PTRS | Align-Di-L | Max |
|---|---|---|---|---|---|---|
| a. (223)(3) | | *! | | | * | * |
| b. (23)(23) | | | *! | * | * | |
| ☞c. (3)(223) | | | | * | | * |
| ☞d. (2223) | | | | * | | * |

b. PP is *sw* (Zhang 1997:316)

| wswS 3333 | PTAS | *33 | Cl | PTRS | Align-Di-L | Max |
|---|---|---|---|---|---|---|
| a. (223)(3) | | *! | | * | * | * |
| b. (23)(23) | | | *! | | * | |
| ☞c. (3)(223) | | | | * | | * |
| ☞d. (2223) | | | | * | | * |

In (11a) where PP is *ws*, Candidate (c) (T3)(**T2T2**T3) and Candidate (d) (**T2T2T2**T3) are optimal outputs. In (11b) where PP is sw, Candidate (c) (T3)(**T2T2**T3) and Candidate (d) (**T2T2T2**T3) are optimal outputs. Notice that the two optimal outputs are identical in the two tableaux. Nevertheless, the union of the winners from the two tableaux does account for the two surface patterns in the sentence *Gou bi wo xiao* 'Dogs are smaller than me', which has a PP containing a pronoun.

In sum, the variant patterns are produced by two inputs with different prosodic strength for PP and two evaluation tableaux. The outputs are combinations of the winners from both tableaux.

**3. Background for modeling variation in OT**

A question we raised earlier was: how do we model variation in OT? There have been different approaches: (i) Partially ordered constraints (Kiparsky 1993, Reynolds 1994, Anttila 1997 *et seq*., Anttila & Cho 1998), (ii) Unranked constraints plus two inputs and two tableaux (Zhang 1997), and (iii) Rank-ordering of EVAL (Coetzee 2006). In what follows, these approaches[2] are briefly introduced.

[2] Probabilistic OT models with numerically valued constraints, such as Stochastic OT and Harmonic Grammar are not discussed here (see Coetzee & Pater in press, for a review and references).

**Partially ordered constraints** (Kiparsky 1993, Reynolds 1994, Anttila 1997 et seq., Anttila & Cho 1998)

In standard OT, a grammar of a language is a total ordering of a ranked set of constraints, typically yielding a single output from each input. However, in a model with partially ordered constraints, a grammar is a partial ordering of the constraint set. Each time when the grammar evaluates a candidate set, one of the rankings consistent with the partial order is randomly chosen. Variation results when some of these rankings select different outputs. The illustration in (12) shows a partial ordering where C is ranked above C1, and C is ranked above C2, but C1 and C2 are crucially unranked.

(12) Partial ordering: C >> C1 and C >> C2

| | C | C1 | C2 |
|---|---|---|---|
| ☞ Candidate (a) | | | * |
| Candidate (b) | | *! | |
| Candidate (c) | *! | | |

| | C | C2 | C1 |
|---|---|---|---|
| Candidate (a) | | *! | |
| ☞ Candidate (b) | | | * |
| Candidate (c) | *! | | |

In (12), since C1 and C2 are crucially unranked (partially ordered) and either C >> C1 >> C2 or C >> C2 >> C1 is consistent with the partial ordering. Both Candidates (a) and (b) in (12) are possible outputs of the grammar.

**Unranked constraints plus two inputs and two tableaux** (Zhang 1997)

Zhang's (1997) analysis with unranked constraints plus two inputs and two tableaux has been presented in § 2.3. The approach is summarized in (13).

(13) Unranked constraints plus two inputs and two tableaux (Zhang 1997)

a. Unranking of some constraints may generate more than one output since some candidates may have a tie under the evaluation.

b. Two inputs are fed into two separate tableaux, potentially producing additional variants.

c. Grammatical outputs are the union of the optimal outputs from the two tableaux.

**Rank-ordering of EVAL** (Coetzee 2006)

In Coetzee's (2006) rank-ordering approach, EVAL not only distinguishes winners from losers, but also imposes a well-formedness ranking-ordering on the whole candidate set. Variation does not arise as a result of variation in grammar (i.e. ranking) as in the previous models; rather, variation depends on how EVAL imposes on a well-formedness rank-ordering on the candidates. In most cases, only the topmost candidate on the rank-ordering is well-formed and grammatical, but under some circumstances, two or more candidates can be well-formed enough to be considered grammatical. Coetzee (2006:338) suggests, "What needs to be added to the grammar is a mechanism that will allow, in some circumstances, more than one of the already generated possible output forms to become actual outputs."

Importantly, there is only one consistent ranking, and added to the evaluation tableau is a critical cut-off line, below which constraint evaluation does not rule out candidates. That is, variation occurs when two or more candidates are not ruled out by higher ranked constraints upon reaching the cut-off. The relative degree of well-formedness of those candidates that pass through the cut-off indicates the relative frequency of the variants. Let us now look at the tableau which illustrates the Rank-ordering Model.

(14) Illustration of the Rank-ordering Model Coetzee (2006:343)

Cut-off

| | C1 | C2 | C3 | C4 |
|---|---|---|---|---|
| ☞ a. Cand-1 | | | | * |
| ☞ b. Cand-2 | | | * | |
| c. Cand-3 | | *! | | |
| d. Cand-4 | *! | | | |

The constraint set is divided into two strata, separated by the critical cut-off. The stratum above the cut-off is the same as classic OT where non-optimal candidates do not surface. The candidates that survive upon reaching the cut-off are grammatical outputs. As we see in (14), Candidates 1 and 2 are winners, i.e. grammatical outputs. Candidates 3 and 4 have fatal violations above the cut-off, hence are ungrammatical. Violations of the constraints in the stratum below the cut-off are not fatal. Variation arises when two or three candidates are disfavored only by constraints below the cut-off, and the extent of constraint violation (i.e. the degree of well-formedness) suggests relative frequency. Candidate 1 and Candidate 2 both have one violation below the cut-off. However, Candidate 1 occurs more frequently than Candidate 2 because it has a higher degree of well-formedness by violating a lower ranked constraint.

## 4. Proposed OT analysis of Mandarin T3S

In this section, we present our proposed OT analysis of Mandarin T3S based on Coetzee's (2006) Model.

### 4.1 Constraints

The constraints, along with their definitions, that we used in our analysis are given in (15).

(15) Constraints

a. ID (T): The tone in the output is faithful to its tone in the input.
b. ID (T, PREP): The tone of a preposition is faithful to its tone in the input.
c. *(33): No adjacent T3's within the same prosodic domain (based on Zhang (1997) with a slight revision).
d. *33: No adjacent T3's, even if they belong to different prosodic domains.
e. *DEG-FT (Degenerate Foot): Monosyllabic feet are banned.
f. BOUND (BOUNDNESS): For every prosodic domain that has more than two syllables, assign a violation mark (based on Chen (2000) with a slight revision).
g. SUBJ-PRED. BOUNDARY: There is a strong subject-predicate boundary. No prosodic domain should go across this boundary.
h. MATCH (PROS, SYN): For each prosodic domain, there must be some syntactic constituent that matches with it. If a prosodic domain does not have a syntactic domain that matches with it, assign a violation mark.

e.g. [Gou [[bi ma] xiao.]]
dog than horse small
'A dog is smaller than a horse.'

| | | | | |
|---|---|---|---|---|
| 3 | 3 | 3 | 3 | UT |
| (3) | (2 | 2 | 3) | → no violation |
| (2 | 2 | 2 | 3) | → no violation |
| (2 | 3) | (2 | 3) | → two violations |

i. WRAP (FC-FC)[3]: Adjacent prosodically weak functional elements in the same PP belong to the same prosodic domain (based on the phonology-syntax interface constraint, WRAP-XP, which prevents XPs from splitting up into multiple p-phrases in Truckenbrodt (1999, 2007)).

[3] A possible motivation for WRAP (Fc, Fc) might be that consistency of directionality of cliticization is preferred. That is, this constraint prevents a situation where the first element cliticizes to the left and the second element to the right.

| | | | | | |
|---|---|---|---|---|---|
| e.g. | [Gou | [[**bi** | **wo]** | xiao.]] | |
| | dog | than | I | small | |
| | ‘A dog is smaller than me.’ | | | | |
| | 3 | 3 | 3 | 3 | UT |
| | (3) | (2 | 2 | 3) | → no violation |
| | (2 | 2 | 2 | 3) | → no violation |
| | *(2 | 3) | (2 | 3) | → violation |

**4.2 The cut-off point and the ranking process**

An essential question to ask is: how should these constraints be ranked and how is the cut-off point determined? Coetzee proceeds by listing *Violation Profiles of Observed Outputs* first (2006:350) and suggests the following procedure of determining where to place the cut-off point. Suppose we have two variant outputs. The input, the two variant outputs, and the constraints each output violates are listed. The constraints that both variant outputs violate are underlined. The underlined constraints cannot distinguish the variants since they are violated by both outputs. Crucially, these constraints must be placed below the cut-off to ensure the two variants do surface. It should be kept in mind that variant outputs must not be disfavored by a constraint above the cut-off. When the frequencies of two variants are available, and we know that one occurs more frequently than the other, we will need some constraint that can distinguish the two. For instance, if *Variant A* that violates *Constraint 1* is higher in frequency than *Variant B* that violates *Constraint 2*, *Constraint 1* has to be ranked lower than *Constraint 2*. By ranking *Constraint 1* lower, a violation of it is less serious, which reflects that *Variant A* is more well-formed, and therefore, occurs more frequently. In addition, in the initial ranking, Coetzee (2006:350) points out that Faithfulness constrains should be ranked low, in accordance with the notion of ranking conservatism (Tesar & Smolensky 1998), unless we have contrary evidence.

Next, we have to ensure that the two variants mentioned above are indeed the best outputs, that is, more well-formed than the rest of the candidates in the competition. According to Coetzee (2006), to ensure that there are no additional grammatical outputs other than the two variant outputs, the rest of the candidates must be eliminated by some constraint above the cut-off. Placing those constraints that are fatally violated by the ungrammatical outputs above the cut-off will guarantee that these output candidates can no longer compete upon reaching the cut-off. That is, for these ungrammatical output candidates, the competition ends at the cut-off point. In this model, constraint ranking is established upon what is observed in the data— what are the variant outputs, and what are not. Then, the constraint ranking is adjusted step by step. We provide selected examples of how we rank our constraints, but omit the whole step-by-step procedure in the adjustment of the ranking argument due to the space limit.

We begin the procedure by listing the constraints that are violated by optimal outputs in five sentences. The violation profiles of the observed outputs are shown in

Table 1. Following Coetzee (2006:350), the constraints shared by the optimal outputs are underlined. This indicates that these constraints can neither distinguish the variants nor eliminate them.

Table 1: Violation profiles of the observed outputs:

| Input | Variants in the Output | Constraints violated |
|---|---|---|
| a. [[*Ma*] [[*hen shao*] *hou*.]] (Zhang 1997:305)<br>horse very rarely roar<br>'Horses very rarely roar.'<br>[[σ] [[σ σ] σ]]<br>3 3 3 3 UT | (3) (223) | <u>BOUND</u>, *Deg-Ft, <u>ID (T)</u> |
| | (2223) | <u>BOUND</u>, <u>ID (T)</u> |
| b. [[*Gou*] [[***bi*** *ma*] *xiao*.]] (Zhang 1997:293)<br>dog than horse small<br>'Dogs are smaller than horses.'<br>[[σ] [[**σ** σ] σ]]<br>**Clitic**<br>3 3 3 3 UT | (3) (223) | ID (T, PREP), BOUND, <u>ID (T)</u>, *Deg-Ft |
| | (2 2 2 3) | ID (T, PREP), BOUND, <u>ID (T)</u> |
| | (2 3) (2 3) | MATCH (PROS, SYN), <u>ID (T)</u> |
| c. [[*Gou*] [[***bi wo***] *xiao*.]] (Zhang 1997:307)<br>dog than I small'<br>Dogs are smaller than I.'<br>[[σ] [[**σ σ**] σ]]<br>**Clitic Clitic**<br>3 3 3 3 UT | (3)(223) | <u>ID (T, PREP)</u>, <u>BOUND</u>, <u>ID (T)</u>, *Deg-Ft |
| | (2 2 2 3) | <u>ID (T, PREP)</u>, <u>BOUND</u>, <u>ID (T)</u> |
| d. [[*Lao Li*] [*zhao xie*.]] (Lin 2007:211)<br>Old Li look for shoes<br>'Old Li looks for shoes.'<br>[[σ σ] [σ σ]]<br>3 3 3 2 UT | (23)(32) | <u>ID (T)</u>, *33 |
| | (2232) | <u>ID (T)</u>, BOUND |
| e. [[*wo*] [*xiang* [*mai bi*.]]] (Lin 2007:215)<br>I want buy pen<br>'I want to buy pens.'<br>[[σ] [σ [σ σ]]]<br>3 3 3 3 UT | (23)(23) | MATCH (PROS, SYN), <u>ID (T)</u> |
| | (2223) | BOUND, <u>ID (T)</u> |

To rank the constraints conservatively, the initial ranking is to place the Markedness constraints over Faithfulness constraints, unless some evidence indicates otherwise. (16) shows the initial ranking.

(16) Initial ranking:

*(33), *33, Match (Pros, Syn), *Deg-Ft, Bound

|

Id (T), Id (T, Prep)

In (17), the arrows indicate from which strata the constraints that are underlined are moved. Adjustment of the ranking is made based on the violation information in Table 1. The constraint *(33) has been moved up to the highest in the ranking because it is never violated. Id (T, Prep) is sometimes violated, and sometimes not, so it is moved to the stratum that has constraints that are not always violated by the observed variants.

(17) Interim ranking

***(33)***

|

Match (Pros, Syn), *Deg-Ft, Bound, **Id (T, Prep )**, *33

|

Id (T)

Interim ranking in (17) is not final and is used here for illustration purposes. Coetzee (2006) approaches his data by gradually ordering the constraints with the evidence of the variants. We follow the same procedure and make adjustments as evidence arises. These steps are critical in identifying exactly where the cut-off point is. In (18), we illustrate the partial ranking of our analysis with a simple sentence from Zhang (1997:295).

(18)

| | | | | |
|---|---|---|---|---|
| [You | [[liang | wan] | mi.]] | |
| there is | two | bowl (CL) | rice | 'There are two bowls of rice.' |
| 3 | 3 | 3 | 3 | UT |
| (3) | (**2** | **2** | 3) | ST1 |
| (**2** | **2** | **2** | 3) | ST2 |

| [[σ][[σσ]σ]] | *(33) | Match (Pros, Syn) | Bound | *33 | Id (T) | *Deg-Ft |
|---|---|---|---|---|---|---|
| ☞ a. (3) (2 2 3) | | | * | | ** | * |
| ☞ b. (2 2  2 3) | | | * | | *** | |
| c. (2 3) (2 3) | | **! | | | ** | |
| d. (2 2 3) (3) | | * | *! | * | ** | * |
| e. (3) (233) | *! | | * | * | * | * |
| f. (33) (23) | *! | ** | | * | * | |

Crucially, ST1 (T3)(**T2T2**T3) in (18) provides evidence that degenerate foot, although disfavored prosodically, is in one of the surface patterns. This indicates that violation of the *Deg-Ft is not serious. It must be ranked below the cut-off. Otherwise, *Candidate (a)* would be ruled out by this constraint before reaching the cut-off, and that would be contradictory to the attested pattern ST1 in (18).

After following Coetzee's procedure, the final ranking reached is shown in (19).

(19) Final ranking argument

*(33), WRAP (Fc, Fc)
|
MATCH (PROS, SYN), BOUND, ID (T, PREP), *33
|
ID (T), *Deg-Ft, Subj.-pred. Boundary

In the next section, we use this final ranking to account for T3S variation in sentences given in (3) and (4).

## 4.3. The analysis

Consider first the sentence without a preposition, analyzed in (20).

(20) Without a preposition:
[*M*a [[*hen* *shao*] *hou*.]]
horse very rarely roar ‘Horses very rarely roar.’

| [σ [[σσ] σ]] | *(33) | WRAP | MATCH (PROS, SYN) | BOUND | ID (T, PREP) | *33 | ID (T) | *DEG-FT | SUBJ.-PRED. BOUNDARY |
|---|---|---|---|---|---|---|---|---|---|
| ☞ a. (3) (2 2 3) | | | | * | | | ** | * | |
| ☞ b. (2 2 2 3) | | | | * | | | *** | | * |
| c. (2 3) (2 3) | | | **! | | | | ** | | * |
| d. (2 2 3)(3) | | | * | *! | | * | ** | * | * |
| e. (2 3) (3 3) | *! | | ** | | | ** | * | | * |

Above the cut-off, the violation of *(33) rules out Candidate (e), MATCH (PROS, SYN) rules out Candidate (c), and both MATCH (PROS, SYN) and BOUND rule out Candidate (d). Candidates (a) and (b), both violating BOUND once, constitute a tie upon reaching the cut-off. Below the cut-off, no violation is fatal. Both Candidates (a) and (b) are grammatical. With respect to relative frequency, Candidate (a) has three violation marks and Candidate (b) has four, which suggests that Candidate (a) is more well-formed and hence occurs more frequently than Candidate (b).

The tableaux in (21) and (22) present our analysis of the sentences with a PP. In (21) where a noun follows the preposition, there are three variants.

(21) With a preposition:
[*Gou* [[*bi* *ma*] *xiao*.]]
dog than horse small 'Dogs are smaller than horses.'

| [σ [[σσ] σ]] | *(33) | WRAP | MATCH (PROS, SYN) | BOUND | ID (T, PREP) | *33 | ID (T) | *DEG-FT | SUBJ.-PRED. BOUNDARY |
|---|---|---|---|---|---|---|---|---|---|
| ☞ a. (3)(2 2 3) | | | | * | * | | ** | * | |
| ☞ b. (2 2 2 3) | | | | * | * | | *** | | * |
| ☞ c. (2 3) (2 3) | | | ** | | | | ** | | * |
| d. (2 2 3) (3) | | | * | * | *! | * | ** | * | * |
| e. (2 3) (3 3) | *! | | ** | | | ** | * | | * |

Above the cut-off, Candidate (e) fatally violates *(33). Candidate (d) violates MATCH (PROS, SYN), BOUND, ID (T, PREP), and *33 and incur 4 violations upon reaching the cut-off and is out of the competition. Candidates (a-c) are tied as they all have two violation marks upon reaching the cut-off. Below the cut-off, no violation is fatal. Candidates (a), (b) and (c) are all grammatical outputs, although they have different numbers of violations (different degrees of well-formedness). Regarding relative frequency, Candidates (a) and (c) both have three violation marks, and Candidate (b) has four. This suggests that Candidates (a) and (c) are more well-formed, and therefore occur more frequently than Candidate (b). In contrast, as analyzed in (22), when a pronoun follows the preposition, there are two variants.

(22) PP with a pronoun
[*Gou* [[*bi* *wo*] *xiao*.]]
dog than I small 'Dogs are smaller than I.'

| [σ [[σσ] σ]] | *(33) | WRAP | MATCH (PROS, SYN) | BOUND | ID (T, PREP) | *33 | ID (T) | *DEG-FT | SUBJ.-PRED. BOUNDARY |
|---|---|---|---|---|---|---|---|---|---|
| ☞ a. (3)(2 2 3) | | | | * | * | | ** | * | |
| ☞ b. (2 2 2 3) | | | | * | * | | *** | | * |
| c. (2 3) (2 3) | | *! | ** | | | | ** | | * |
| d. (2 2 3)(3) | | | * | * | *! | * | ** | * | * |
| e. (2 3) (3 3) | *! | * | ** | | | ** | * | | * |

Above the cut-off, Candidate (c) fatally violates the WRAP constraint since the two elements of the PP are not parsed into the same prosodic domain. Candidate (e) violates both *(33) and WRAP in the top tier of the constraints, and is ruled out. Candidate (d) violates MATCH (PROS, SYN), BOUND, ID (T, PREP), and *33 and incurs 4 violations and is ruled out upon reaching the cut-off. Both Candidate (a) and Candidate (b) incur two violation marks upon reaching the cut-off. They are both grammatical outputs. Below the cut-off, Candidate (a) has three violation marks and Candidate (b) has four, which suggests that Candidate (a) is more well-formed and hence occurs more frequently than Candidate (b).

### 4.4. Summary

In sum, the one-to-many in the input-output mapping in T3S is a challenging aspect in T3S research and the variation is further complicated by the prepositions and pronouns. There has been little progress in analyzing these data in the OT framework. By adopting Coetzee’s (2006) OT variation model, our proposed re-analysis of the T3S variation in the examples with prepositions and pronouns have the following merits: (i) This is a one-input and one-tableau analysis and multiple surface patterns are produced in one step, (ii) there is no need to assume that PP is unspecified for prosodic strength, and (iii) the relative frequency of the multiple surface forms is indicated, against which empirical data can be checked.

## 5. Concluding remarks

We start out with a few simple examples to illustrate that by adopting Coetzee's (2006) OT variation model, variation in Mandarin T3S can be analyzed with one input, one ranking, and one evaluation tableau. The encoding of relative well-formedness or relative frequency of the variants allows the proposal to provide the basis for future sociolinguistic and experimental studies. The notion of different degrees of well-formedness is also applicable to other domains. Preliminary results from one of our experiments show that children typically make the 'better' kinds of errors and the rank-ordering model provides a way to capture children's T3S error patterns.

Future studies will need to expand to additional examples with various syntactic structures, collect empirical data, and explore and compare several OT variation models. Careful sociolinguistic studies of the variation in Mandarin T3S are necessary to better understand the distribution of the surface variants, which can then be modeled formally. Lastly, little is known about how children acquire T3S in phrases and sentences, let alone the variation patterns through the developmental stages. It would be interesting to find out whether or not or the extent to which children produce the same variants attested in adults, and whether or not the relative frequency of children's patterns is similar to those of adults’.

## REFERENCES

ANTTILA, ARTO. 1997. Deriving variation from grammar. *Variation, change, and phonological theory*, ed. by Frans Hinskens, Roeland Van Hout, and W. Leo Wetzels, 35–68. Amsterdam: John Benjamins.

ANTTILA, ARTO and YOUNG-MEE YU CHO. 1998. Variation and change in optimality theory. *Lingua* 104.31-56.

CHEN, MATTHEW Y. 2000. *Tone sandhi: Patterns across Chinese dialects.* Cambridge: Cambridge University Press.

CHENG, CHIN-CHUAN. 1973. *A synchronic phonology of Mandarin Chinese.* Mouton: The Hague.

CINQUE, GUGLIELMO. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24.239-97.

COETZEE, ANDRIES W. 2006. Variation as accessing 'non-optimal'candidates. *Phonology* 23.337-85.

COETZEE, ANDRIES W. and JOE PATER. in press. The place of variation in phonological theory. *The handbook of phonological theory*, ed. by John A. Goldsmith, Jason Riggle, and Alan C. L. Yu. Oxford: Wiley-Blackwell.

DUANMU, SAN. 2000/2007. *The Phonology of Standard Chinese*. 1st edn./2nd edn. Oxford: Oxford University Press.

KIPARSKY, PAUL. 1993. An OT perspective on phonological variation. Handout from *Rutgers Optimality Workshop 1993*, also presented at *NWAV 23* 1994, Stanford University. On-line: http://www.stanford.edu/~kiparsky/ Papers/nwave94.pdf.

LIN, YEN-HWEI. 2007. *The sounds of Chinese*. Cambridge: Cambridge University Press.

REYNOLDS, WILLIAM T. 1994. *Variation and phonological theory*. Philadelphia, PA: University of Pennsylvania dissertation.

SHIH, CHILIN. 1986. *The prosodic domain of tone sandhi in Chinese.* La Jolla, CA: University of California San Diego dissertation.

SHIH, CHILIN. 1997. Mandarin third tone sandhi and prosodic structure. *Studies in Chinese phonology*, ed. by Norval Smith and Jialing Wang, 81-123. Berlin: Mouton de Gruyter.

TESAR, BRUCE and PAUL SMOLENSKY. 1998. Learnability in optimality theory. *Linguistic Inquiry* 29.229-68.

TRUCKENBRODT, HUBERT. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30.219-55.

TRUCKENBRODT, HUBERT. 2007. The syntax-phonology interface. *The Cambridge handbook of phonology*, ed. by Paul de Lacy, 435-56. Cambridge: Cambridge University Press.

ZHANG, NING. 1997. The avoidance of the third tone sandhi in Mandarin Chinese. *Journal of East Asian Linguistics* 6.293-338.