An Analysis of the Xiamen Tone Circle

Guillaume Thomas
Massachusetts Institute of Technology

1. Introduction

1.1. The Xiamen Tone Circle

Xiamen[1][2] is a Southern Min language that exhibits a complex system of tone sandhi: the tone that is associated with a given word (most words being monosyllabic) varies depending on the prosodic/syntactic context in which the word is used. A monosyllabic word that appears in isolation or that appears in a prominent position in a tone group bears what has been called its citation tone. When the word is used in a non prominent position in a tone group, it bears a different tone called its sandhi tone.

The definition of tone groups is itself a matter of ongoing research. Suffice it to say that tone groups are prosodic groups of words whose most prominent position is the syllable that is aligned with the right edge of the group, cf. Chen (1987) and Hsieh (2004). Tone sandhi in Southern Min languages generally takes place in non final (ie. non prominent) positions in the tone group.

In Xiamen, citation tones are mapped to sandhi tones in a systematic way. Sandhi tones form a proper subset of the set of citation tones, the global mapping from citation tones to sandhi tones forming a circle, as illustrated by the following diagram:

$$\begin{tikzcd} & & 21 \arrow[dr] & \\ 24 \arrow[r] & 22 \arrow[ur] & & 53 \arrow[dl] \\ & & 44 \arrow[ul] & \end{tikzcd}$$

The following table from Chen (1987) illustrates the phenomenon with a few words in citation/sandhi forms:

(1)	Citation Form	Translation	Sandhi Form	Translation
	we-24	‘shoe’	we-22 tua-21	‘shoe laces’
	wi-22	‘stomach’	wi-21 pih-22	‘stomach ailment’
	tsu-21	‘house’	tsu-53 ting-53	‘roof top’
	hai-53	‘ocean’	hai-44 kih-24	‘ocean-front’
	pang-44	‘fragrant’	pang-22 tsui-53	‘fragrant water’

1.2. A problem for classical OT

The Xiamen tone circle is one of the non-computable functions in OT identified by Moreton (1999). As such, it is not analyzable in classical OT, i.e. using only Markedness and Faithfulness constraints and

[1]Many thanks to Adam Albright, Donca Steriade and Michael Kenstowicz for their time and their valuable comments. I am also indebted to Adam for help with Perl. Thanks to the Phonology Circle at MIT and especially to Edward Flemming for comments and criticisms on a previous version of this handout. All errors are mine.

[2]I will ignore checked tones, i.e. tones that appear on syllables which are closed with an obstruent.

 Proceedings of the 27th West Coast Conference on Formal Linguistics, ed. Natasha Abner and Jason Bishop, 422-430. Somerville, MA: Cascadilla Proceedings Project.

a standard conception of GEN and EVAL. The problem that is raised by the tone circle is that of any circular mapping in OT. We can illustrate it by a simple circular mapping between two forms, A and B:

(2) { /A/ → [B], /B/ → [A] }

Let us assume the markedness constraints *A and *B, and the faithfulness constraints FAITH_A and FAITH_B. Motivating /A/ → [B] would require ranking *A over *B, and *A over FAITH_A. Motivating /B/ → [A] would require ranking *B over *A, and *B over FAITH_B. A classical OT analysis thus faces a ranking contradiction: *A >> *B and *B >> *A.

This paper proposes a simple analysis of the Xiamen tone circle in an extended version of OT, using a new family of linear-faithfulness constraints.

2. Two fundamental assumptions: Structure Preservation and Systemic Evaluation

2.1. Structure Preservation

As we have already seen, the set of sandhi tones is a subset of the set of citation tones – a phenomenon we might attribute to structure preservation, cf. Kiparsky (1985). In the current analysis this can be expressed through the assumption of a number of undominated markedness constraints which define the inventory of possible tones in Xiamen, corresponding to the surface citation tones. Since the mapping from citation to sandhi tones has to satisfy these undominated constraints, the set of sandhi tones can only be a subset of the set of citation tones. I will not try to formulate the constraints that are responsible for structure preservation. This set should include for example constraints preventing the realization of tones with more than two targets, or tones whose targets differ by more than two Chao-units, among other constraints.

All the way through, I will assume that tones are represented as ordered pairs of Chao-units from 1 to 5 representing the pitch of the first and the second target of the tones respectively. The set of possible citation/sandhi tones of Xiamen is {21, 22, 24, 44, 53}. This representation of tones has been chosen over register and shape representation (cf. Yip (2002)) because it simplifies the analysis, which does not make any direct reference to the register of the tones.

2.2. Systemic Evaluation

I will assume an OT grammar in which the candidates for evaluation are mappings from the set of citation tones to the set of sandhi tones. A GEN-like function is assumed to take the set of attested citation tones as its argument and to generate from it mappings from the set of possible citation tones to the set of possible sandhi tones. More precisely, every candidate is a mapping of every tone in the set {21, 22, 24, 44, 53} to a single tone in the same set. Ideally, candidates should include any mapping from some set of tones to some other set of tones, but since we assume that high ranking markedness constraints (structure preservation) define the sets of possible citation and sandhi tones as the same set {21, 22, 24, 44, 53}, we can restrict our attention to mappings from this set to itself. Both one to one and many to one mappings are elligible, since we do not want to rule out *a priori* the possibility of a merger in the citation to sandhi mapping. We require that every candidate map every possible citation tone (i.e. every member of the set {21, 22, 24, 44, 53}) to a possible sandhi tone, but not the other way round.

In the limits of these assumptions, we will have to consider 3,125 candidates, the number of permutations with repetition of the sequence of possible tones <21, 22, 24, 44, 53>.

The faithfulness constraints that we will consider will be defined over relations between citation tones and their corresponding sandhi tones in a given candidate mapping. In that sense, they should be thought of as Output-Output constraints.

The following diagrams present one of the candidates in two different notations:

(3) {/24 22 21 53 44/ → [22 22 21 21 44]} or:

$$\begin{tikzcd}
 & & 44 & \\
24 \arrow[r] & 22 & & 53 \arrow[dl] \\
 & & 21 &
\end{tikzcd}$$

I will call this approach to the generation and evalution of candidates ‘systemic evaluation’, cf. Flemming (2002). See Barrie (2006) for a similar approach to the Xiamen tone circle.

3. Triggering the Circle

Looking at the tone circle as a phenomenon of alternation between citation tones and sandhi tones, we must ask the two following questions. Why is there an alternation at all? Why is the alternation structured as a chain shift?

3.1. Motivating an alternation: *RISE

A striking property of the circle suggests the answer to the first question. All the tones that are present in citation are present in sandhi, except one: 24. This tone appears to be the only rising tone of the whole system. We might assume that a high ranking markedness constraint against rising tones *RISE is responsible for its loss in sandhi:

(4) *RISE: One violation for this constraint is incurred by every sandhi tone that has a rising contour

I will assume that *RISE is defined over sandhi tones only. This is a simplification that is justified by the fact that, because citation tones occur in prominent prosodic positions, they are more faithful to their input than sandhi tones are to their citation correspondent. To be more precise, it has been shown (cf. ao. Lin (1988) for the closely related dialect of mainstream Taiwanese) that syllables in citation have a longer duration than syllables in sandhi, which explains the presence of rising tones in citations and their absence in sandhi, in accordance with the correlation between duration and contour-bearing ability that is argued for in Zhang (2001). I assume that an IO positional faithfulness constraint IDENTSHAPE(PROMINENT) dominates *RISE and permits the rising tones to surface as such in citation. But since we are here concerned with OO relations between the fixed set of attested citation tones and their sandhi correspondent, we can overlook this positional faithfulness constraint and consider that *RISE is defined over sandhi tones only. The use of a *RISE constraint in the analysis of Southern-Min tonal chain shifts (especially mainstream and coastal Taiwanese) has already been argued for in Hsieh (2004).

3.2. Triggering a chain shift: *MERGE *and* DIFFER

Whereas a high-ranking *RISE can be used to trigger a mapping from a rising tone in citation to a non rising tone in sandhi, it will not trigger a chain shift. We want to exclude candidates such as the follwing one:

(5)

$$\begin{tikzcd}
 & & 21 & \\
24 \arrow[r] & 22 & & 53 \\
 & & 44 &
\end{tikzcd}$$

I assume that an anti-faithfulness constraint (cf. Alderete (2001)) DIFFER will prevent other citation tones from being mapped to themselves in sandhi:

(6) DIFFER: One violation for this constraint is incurred by every mapping of a citation tone to a sandhi tone to which it is identical.

Since all the citations tones of Xiamen are mapped to a sandhi tone other than themselves, we can assume that DIFFER is undominated.

Note that although Anti-faithfulness and Anti-merger constraints might be thought of as alternative ways to trigger a chain-shift (in addition with a high ranked markedness constraint), DIFFER will not prevent various citation tones to be mapped to the same sandhi tones. There is indeed one case of merging of citation tones in sandhi in Xiamen:

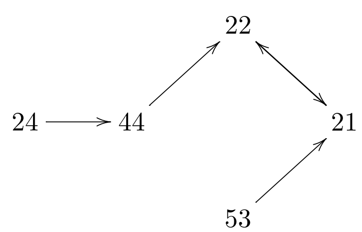
(7) { /24/ → [22], /21/ → [22] }

This case of merger is of course inevitable if 24 is to be excluded from the set of sandhi tone by an undominated *RISE, and the system of tone sandhi obeys structure preservation. But this is the only case of merger in the attested chain-shift. In order to rule out candidates with more cases of merger, we can introduce an anti-merger constraint (cf. Padgett (2003), Łubowicz (2003), Hsieh (2004)), *MERGE:

(8) *MERGE: One violation for this constraint is incurred by every pair of different citation tones that are mapped to the same sandhi tone.

This constraint will allow us to exclude candidates like the folowing one, that are excluded by *RISE and DIFFER alone:

(9)



Violations:

*RISE: 0 DIFFER: 0 *MERGE: 2

3.2.1. What we have done so far

Since *RISE in conjunction with structure preservation requires a merger of two citation tones in one sandhi tone, we have to rank *RISE over *MERGE. There does not seem to be any conflict between DIFFER and *RISE and *MERGE, and we might rank DIFFER over *MERGE by default. Thus we get the following grammar:

(10) *RISE, DIFFER $>>$ *MERGE

It can be show that all the candidates that incur no violation of *RISE and DIFFER and only one violation of *MERGE obey one of the two following structural descriptions:

(11)

$$\begin{tikzcd} 24 \arrow[r] & A \arrow[r] & B \arrow[r] & C \arrow[r] & D \arrow[lll, dotted, bend right, "or"'] \arrow[ll, dotted, bend right, "or"'] \arrow[l, dotted, bend left, "or"] \end{tikzcd} \qquad 72$$

$$\begin{tikzcd} 24 \arrow[r] & A \arrow[r] & B \arrow[l, bend left] & C \arrow[r] & D \arrow[l, bend left] \end{tikzcd} \qquad 12$$

where A≠B≠C≠D and $A, B, C, D \in \{21, 22, 44, 53\}$ and dotted arrows represent alternative mappings

There are only 84 such mappings among the 3,125 possible candidates. Thus the simple grammar in (10) filters all but 84 candidates.

4. Structuring the chain shift: DRAT, IDENTPITCH and IDENTSHAPE

We will now work at extending this grammar with constraints that can narrow down the 84 surviving candidates to the attested winner in a principled way. I propose three such constraints: DRAT, IDENTPITCH and IDENTSHAPE.

4.1. DRAT

In the attested mapping, three citation tones are mapped to a sandhi tone that has lower targets or equally high targets:

(12) /24/ → [22], /22/ → [21], /44/ → [22]

And two tones are mapped to a sandhi tone that has at least one higher target than its corresponding target in citation:

(13) /21/ → [53], /53/ → [44]

Hence, it seems that this system favors candidates that minimize the number of mapping between a citation tone and a sandhi tone that is globally higher in pitch. By a ‘globally higher pitch’ I mean to refer to mappings in which at least one target of a sandhi tone is higher than the corresponding target in citation. Let us clarify this notion with the help of the following definitions:

(14) Correspondance between targets:

- The first target of a sandhi tone corresponds to the first target of a citation tone that is mapped to it. The second target of a sandhi tone corresponds to the second target of a citation tone that is mapped to it. A target of a sandhi tone cannot correspond to any other targets than those defined above.
- /21/ → [53]: 2 corresponds to 5, 1 corresponds to 3

(15) Target Raising:

- A target of a sandhi tone has been raised if and only if it is higher than the corresponding target of the citation tone that is mapped to it.
- /24/ → [53]: 2 is raised, 4 is not

The constraint DRAT captures the general dispreference against pitch raising in the citation to sandhi mapping:

(16) DRAT (Don’t Raise Any Target):

- DRAT A sandhi tone incurs one violation for this constraint if and only if at least one of its targets has been raised.

The following ranking filters all but 24 of the 3,125 candidates:

(17) *RISE, DIFFER >> *MERGE >> DRAT

4.2. IDENTPITCH

IDENTPITCH penalizes the distance between a citation tone and its sandhi tone, measured in Chao units that separate corresponding targets:

(18) ID(PITCH): One violation for this constraint is incurred by a sandhi tone for every Chao unit that separates each of its targets from the corresponding target in citation.

As an illustration, { /24/ → [21], /22/ → [53] } incurs 8 violations of ID(PITCH). The following ranking is adopted:

(19) *RISE, DIFFER >> *MERGE >> DRAT >> ID(PITCH)

This grammar filters al but two of the 3,125 candidates.

4.3. Don't change your shape

The following diagram presents the two remaining candidates:

(20)

$$\begin{tikzcd}
 & & 21 \arrow[dr] & \\
24 \arrow[r] & 22 \arrow[ur] & & 53 \arrow[dl] \\
 & & 44 \arrow[ul] &
\end{tikzcd}$$

DRAT: 2 violations
Total number of Chao units raised: 6
ID(PITCH): 14 violations

$$\begin{tikzcd}
 & & 21 \arrow[dr] & \\
24 \arrow[r] & 22 \arrow[ur] & & 44 \arrow[dl] \\
 & & 53 \arrow[ul] &
\end{tikzcd}$$

DRAT: 2 violations
Total number of Chao units raised: 6
ID(PITCH): 14 violations

We note that in the winning candidate, only three citation tones are mapped to a tone with a different shape.

(21) /24/ → [22], /22/ → [21], /53/ → [44]

However, in the loosing candidate, all five citation tones are mapped to a tone with a different shape. A faithfulness constraint that penalizes modification of shape will pick out the winner:

(22) IDSHAPE: One violation for this constraint is incurred by every sandhi tone whose shape differs from the shape of the citation tone that is mapped to it.

The three attested shapes in the Xiamen tonal system are level, rise and fall. I consider that tone shape is not included as a feature in the representation of tones, but is reduced to the relative height of the two targets of each tone.

5. Final Ranking

The grammar consisting of the following ranking of constraints and the system of generation and evaluation of candidates described in 2.2 has been implemented and successfully tested with a Perl script and OTSoft (Hayes et al. (2003)):

(23) DIFFER, *RISE >> *MERGE >> DRAT >> ID(PITCH) >> IDENTSHAPE

The Perl script (cf. appendix) generates all the 3,125 candidates and evaluates their violations of the five constraints that we have introduced. It outputs a tabulated text file that can be pasted and formated in an Excel spreadsheet and evaluated with OTSoft.

6. Linear Faithfulness

There are several aspects of the present proposal that are non classical. Beyond the general architecture of the OT framework we used, the constraint DRAT strikes us as combining elements of a faithfulness constraints (do not change your pitch) and of a markedness constraint (only high pitch is penalized – relative to the initial pitch). I suggest that this constraint belongs to a more general family of Linear Faithfulness constraint, that we might define as follows.

(24) Linear Feature:

A feature ϕ is a linear feature if and only if there is an order relation R on the set V of its possible values, such that R is transitive, antisymmetric and total.

(25) ϕ-Mapping:

Given a feature ϕ, a mapping /A/ $\longrightarrow$ [B] from an input A to an output B is a ϕ-mapping if and only if A and B have possibly different non null values for ϕ, noted respectively ϕ_A and ϕ_B.

(26) Linear Faithfulness Constraint:

C^F is a linear faithfulness constraint if and only if F is a linear feature and there is a precedence relation $\leq$ on the set of its possible values, such that, for any F-mapping /A/ $\longrightarrow$ [B]:
/A/ $\longrightarrow$ [B] violates C^F if and only if $F_A \leq F_B$ and $F_A \neq F_B$

The introduction of such a family of constraints is not strickly speaking a new move in OT. The combination of markedness and faithfulness is already present in the targetted faithfulness constraints of Wilson (2000). Linear Faithfulness constraints are basically faithfulnes constraints that do not penalize just any change of feature value, but a change that is directed on a scale of values. As such, linear faithfulness constraints are another example of the use of scales in OT, a possibility explored in works such as Gnanadesikan (1997) or Mortensen (2004).

Note that conjoining a non gradual faithfulness constraint IDENTPITCH with a markedness constraint *MARKEDPITCH would not yield the same results as our constraint DRAT. Indeed, IDENTPITCH&*MARKEDPITCH would penalize both raising and lowering of a target to a marked (presumably, high) pitch in sandhi. On the contrary, DRAT penalizes only a directed violation of faithfulness to pitch.

It should also be noted that linear faithfulness constraints are more constrained than markedness and faithfulness conjunction. Whereas constraint conjunction allow for any combination of faithfulness and markedness, the use of linear faithfulness constraints entail that markedness (here the penalization of relatively high pitch values) be defined on the set of possible values of the feature with respect to which the faithfulness of the output is being evaluated.

7. Conclusion

We have proposed a rigorous analysis of the Xiamen tone circle using extensions of OT that have been already introduced in the treatment of other phenomena: systemic evaluation (cf. Flemming (2002), Łubowicz (2003) for similar ideas), Anti-Faithfulness and Anti-Merger constraints, and use of scales of feature-values in the form of a Linear Faithfulness constraint.

To our knowledge, this is the first phonological analysis of the Xiamen tone circle in OT that has been demonstrated to derive the attested tonal chain shift – cf. script in appendix. This is a progress with respect to the contrast preservation approach of Barrie (2006), who did not offer any procedure to test the validity of his analysis.

The set of constraints that we have introduced would have to be tested in the analysis of other tonal chain shifts of the same linguistic group, to evaluate its cross-linguistic relevance.

Appendix

```
#/usr/bin/perl

print "\t\t\tIdentShape\tIdentPitch\tDRAT\t*Merge\tDiffer\t*Rise\n";
print "\t\t\tIdentShape\tIdentPitch\tDRAT\t*Merge\tDiffer\t*Rise\n";

@tones = (24,22,21,53,44);
@shape = ("rise","level","fall","fall","level");
@target1 = (2,2,2,5,4);
@target2 = (4,2,1,3,4);

print "@tones\n\n";

for (my $t1 = 0; $t1 < scalar @tones; $t1++) {
for (my $t2 = 0; $t2 < scalar @tones; $t2++) {
for (my $t3 = 0; $t3 < scalar @tones; $t3++) {
for (my $t4 = 0; $t4 < scalar @tones; $t4++) {
for (my $t5 = 0; $t5 < scalar @tones; $t5++) {
@outputs = ($t1,$t2,$t3,$t4,$t5);

# for each constraint, define a variable that
# store its violations, start at 0 and increment

$merge = 0;
$rise = 0;
$ident_shape = 0;
$ident_pitch = 0;
$DRAT = 0;
$Differ = 0;

for (my $count = 0; $count < 5; $count++) {
# *Rise
if ($shape[$outputs[$count]] eq "rise") { $rise++; }

# Faithfulness Constraints

# IdentShape

unless ($shape[$count] eq $shape[$outputs[$count]])
{ $ident_shape++; }

# IdentPitch

$ident_pitch1 = 0; $ident_pitch2 = 0;

unless ($target1[$count] == $target1[$outputs[$count]])
{$ident_pitch1 = (abs($target1[$count] -
$target1[$outputs[$count]])); }

unless ($target2[$count] == $target2[$outputs[$count]])
{$ident_pitch2 = (abs($target2[$count] -
$target2[$outputs[$count]])); }
```

```
$ident_pitch += ($ident_pitch1 + $ident_pitch2);

# Don’t Raise Any Target Unary (DRAT)

if (($target1[$count] < $target1[$outputs[$count]]) ||
($target2[$count]<$target2[$outputs[$count]])){$DRAT++;}

# Anti-Faithfulness

if ($tones[$count] == $tones[$outputs[$count]])
{ $Differ++; }

# the inner loop for assessing violations of constraints that involve
# pairwise evaluation of candidates

for (my $count2 = $count+1; $count2 < 5; $count2++) {

# *Merge

if ($outputs[$count] == $outputs[$count2]) {$merge++;}}}

# print the set of output tones

print"\t";

for (my $count = 0; $count < scalar(@outputs); $count++)
{ print "$tones[$outputs[$count]] "; }

# print the corresponding violationss.

print "\t\t$ident_shape\t$ident_pitch\t$DRAT\t$merge
\t$Differ\t$rise\t";
print"\n"; } } } } }
```

References

Alderete, John (2001). *Morphologically Governed Accent in Optimality Theory*. Outstanding Dissertations in Linguistics, Routledge.

Barrie, Michael (2006). Tone Circle and Contrast Preservation. *Linguistic Inquiry* 37:1.

Chen, Matthew Y. (1987). The Syntax of Xiamen Tone Sandhi. *Phonology* 4, pp. 109–149.

Flemming, Edward (2002). *Auditory Representations in Phonology*. Outstanding Dissertations in Linguistics, Routledge.

Gnanadesikan, Amalia (1997). *Phonology with Ternary Scales*. Ph.D. thesis, UMass Amherst.

Hayes, Bruce, Bruce Tesar & Kie Zuraw (2003). Otsoft 2.1 software package. Http://www.linguistics.ucla.edu/people/hayes/otsoft/.

Hsieh, Feng Fan (2004). On the proper treatment of tonal chain shifts. MIT. Generals Paper.

Kiparsky, Paul (1985). Some consequences of lexical phonology. *Phonology Yearbook* 2.

Lin, Hwei-Bing (1988). *Contextual Stability of Taiwanese Tones*. Ph.D. thesis, The University of Connecticut.

Łubowicz, Anna (2003). *Contrast Preservation in Phonological Mappings*. Ph.D. thesis, UMass Amherst.

Moreton, Elliot (1999). Non-Computable Functions in Optimality Theory. *ROA 36* .

Mortensen, David (2004). Abstract scales in phonology.

Padgett, Jaye (2003). Systemic contrast and catalan rhotics. Ms. ROA 574-0203.

Wilson, Colin (2000). *Targeted Constraints: An approach to contextual neutralization in Optimality Theory*. Ph.D. thesis, John Hopkins University.

Yip, Moira (2002). *Tone*. CUP.

Zhang, Jie (2001). *The effects of duration and sonority on contour tone distribution*. Routledge.

Proceedings of the 27th West Coast Conference on Formal Linguistics

edited by Natasha Abner and Jason Bishop

Cascadilla Proceedings Project Somerville, MA 2008

Copyright information

ISBN 978-1-57473-428-7 library binding

Ordering information

Web access and citation information