Hai-Lu Hakka Shang Tone Sandhi in Optimality Theory Framework*

Chang, Tsai-fang

1. INTRODUCTION

Hai-lu Hakka, a dialect of Hakka language, is spoken in most of the towns in Hsin-chu in Taiwan.[1] It is called so because it is the main dialect of the two Hakka counties of Canton province in China, i.e., Hai-feng and Lu-feng. Hai-lu Hakka is a tone language with seven tone systems. The seven tones are presented as below in (1).

(1) The Seven Tone in Hai-lu Hakka[2]

Tone Inventory	**Represented Tonal Value**	**Example**
A. Yin-ping Tone	HM (high-mid falling tone)	*ko* 'tall'
B. Shang Tone	LM (low-mid raising tone)	*cho* 'grass'
C. Yin-qu Tone	LL (low-low falling tone)	*sin* 'letter'
D. Yin-ru Tone	H (high checked tone)	*kiok* 'foot'
E. Yang-ping Tone	HH (high tone)	*tiam* 'sweet'
F. Yang-qu Tone	MM (mid tone)	*lu* 'road'
G. Yang-ru Tone	M (mid checked tone)	*phak* 'white'

Among the seven tones, Shang Tone and Yin-ru Tone display Tone Sandhi phenomena. The focus of this paper is restricted in the Shang Tone sandhi phenomenon, and the Yin-ru Tone sandhi phenomenon is not discussed in the present paper.

2. SHANG TONE SANDHI PHENOMENON

2.1 Base Tone and Sandhi Tone

Approximately speaking, Shang Tone words preserve the base low-mid raising tone LM[3] when they appear as single words or the last words of phrases. On the other hand, when they appear in the other positions in phrases, they change from the LM tone to the sandhi tone, that is, a low smooth tone.[4] The data in (2) represent this fact.

---

*Thanks to Professor Yuchao E. Hsiao for inspiring me the idea of this paper, any mistakes are, of course, mine.

1. The data presented in this paper is used commonly in Chu-tong town.
2. For more details, see Gu eds. (1998).
3. For ease of discussion, Shang Tone is represented as LM henceforth in this paper.
4. For ease of discussion, the sandhi tone for Shang Tone is represented as LL henceforth in this paper.

(2)

Example	Meaning	Base Tone	Sandhi Tone
(a)*ciu kui*	'a tippler'	LM LM	LL LM
(b)*fu kua*	'a balsam pear'	LM HM	LL HM
(c)*chian sui ciang*	'a shallow well'	LM LM LM	LL LL LM
(d)*fai fu kua*	'a bad balsam pear'	LM LM HM	LL LL HM
(e)*da lo fu*	'to hit the tiger'	LM LM LM	LL LL LM
(f)*mai shui ko*	'to buy fruit'	HM LM LM	HM LL LM

2.2 Derivational Rule Description

From the data in (2) above, the general rule that describes Shang Tone sandhi phenomenon can be characterized as below in (3).

(3) Shang Tone Sandhi Rule:

LM → LL / __T (LM = base tone; LL = sandhi tone; T = any tones)

However, this rule wrongly predicts that the tone of the Shang Tone words in (4) and (5) will change to the sandhi tone LL.

(4) Wrongly Predicted Shang Tone Words (in interphrasal or sentential level)

Example	Meaning	Base Tone	Predicted Tone	Correct Tone
(a)*fo thai*	'very angry'	LM MM	*LL MM	LM MM
(b)*sui kun*	'the water boils'	LM LM HM	*LL LL HM	LM LM HM
(c)*siu chien*	'the hands are kittenish'	LM MM	*LL MM	LM MM

(5) Wrongly Predicted Shang Tone Words (with numerals and classifiers)

Example	Meaning	Base Tone	Predicted Tone	Correct Tone
(a)*liong thung sui*	'two buckets of water'	LM LM LM	*LL LL LM	LM LM LM
(b)*kiu von phon*	'nine bowls of rice'	LM LM MM	*LL LL MM	LM LM MM
(c)*mai pun su*	'to buy one book'	HM LM HH	*HM LL HH	HM LM HH
(d)*shau thung cio*	'to lack a bucket of wine'	LM LM LM	*LL LL LM	LL LM LM

Although phrases in (4) are often used as adjectives in Hai-lu Hakka, they can be considered as whole sentences. Phrases in (4) are constructed by a subject noun and a predicate, thus, they are not like the phrases in (2) and (5). There are boundaries between the subjects and the predicates, so the Shang Tone words in (4) do not alter their tone. The derivational rule must be restricted in the domain of a single XP, in other words, there is a tone sandhi domain boundary at the right edge of every XP. The rewritten rule is shown in (6) below.

(6) Shang Tone Sandhi Rule: **(applied in single XP domain)**

LM → LL /__T (LM = base tone; LL = sandhi tone; T = any tones)

The rule in (6) can successfully predict the correct tone of the Shang Tone words in (5). The derivational steps are shown in (7) below.

(7) Derivational Steps of Shang Tone Words

fo thai	'very angry'	
Base Tone	[LM] [LM]	
Tone Sandhi Rule	NA[5]	first (word) level
	[LM]# [LM]#	output of first level
Tone Sandhi Rule	NA[6]	second (phrasal) level
	[LM LM]	output of second level

It is obvious that Shang Tone numerals and classifiers appearing in (5) do not change their base tone to the sandhi tone. Phrases with classifiers can be regarded as an XP, that is, a domain which the tone sandhi rule can be applied in. This fact will be discussed later in part 5. However, the derivational Shang Tone sandhi rule is not appropriate to account for the Shang Tone numerals. In Hai-lu Hakka, if Shang Tone words surface as numerals, they do not display Tone Sandhi phenomenon. However, if they do not have the syntactic identity as numerals, they are subject to the Shang Tone sandhi rule.[7] To account for this fact, an additional condition must be added to the Shang Tone sandhi rule. The Shang Tone sandhi rule with the additional condition can be described as below in (8).

(8) Shang Tone Sandhi Rule: **(applied in single XP domain and not applied to numerals)**

LM → LL /__T (LM = base tone; LL = sandhi tone; T = any tones)

The application of rule (8) can be accounted as below in (9).

(9) Derivational Steps of Shang Tone Numerals

---

5. NA stands for the situation in which the rule is not applied. The Shang Tone Sandhi rule is not applied here in the first level because the words *fo* 'fire' and *thai* 'big' surface with their base tone individually.

6. The Shang Tone Sandhi rule is not applied here because both the Shang Tone words *fo* 'fire' and *thai* 'big' are the final words of their local phrases.

7. For example, the word *kiu* in *kiu pak* which stands for a person's name is pronounced as the sandhi tone LL.

liong thung sui	'two buckets of water'	
Base Tone	LM LM LM	
Tone Sandhi Rule	NA	first (word) level
	[LM LM]# LM#	output of first level
Tone Sandhi Rule	NA[8]	second (phrasal) level
	[LM LM LM]	output of second level

The Shang Tone Sandhi rule with the additional conditions above can predict the sandhi ton for the Shang Tone words in phrases in several steps. In the next section, an Optimality approac will be introduced.

3. DIFFERENT SYNTACTIC IDENTITIES OF MONOSYLLABIC ADJECTIVES

Shih (1986), Hong (1987) and Hsiao (2000 a) observe that in Mandarin and Taiwanese, monosyllabic adjectives appearing in front of nouns as modifiers are likely to become subject t lexicalization; therefore, they can only be treated as syntactic categories, i.e., As, rather tha maximal projections, i.e., APs. This observation also holds true for Hai-lu Hakka monosyllabi adjectives. The data in (10) show the Tone Sandhi phenomenon of Shang Tone monosyllabi adjectives.

(10)	**Example**	**Meaning**	**Base Tone**	**Sandhi Tone**
	(a)*fu ciu*	'bitter liquor'	LM LM	LL LM
	(b)*ka fa*	'an artificial flower'	LM HM	LL HM
	(c)*si lo su*	'a dead rat'	LM LM LM	LL LL LM
	(d)*fai sip kuan*	'a bad habit'	LM M LL	LL M LL

However, if the monosyllabic adjectives are appearing with an adjacent suffix -kai[9], they do no change their base tone. In other words, they are not subject to lexicalization when they are adde with a suffix -kai, and this situation is shown in (11) below.

(11)	**Example**	**Meaning**	**Base Tone**	**Sandhi Tone**
	(a)*fu kai ciu*	'bitter liquor'	LM LL LM	no sandhi
	(b)*ka kai fa*	'an artificial flower'	LM LL HM	no sandhi

---

8. The Shang Tone sandhi rule is not applied here because the word *thung* 'bucket' is the final word of the classifier phrase *liong thung* 'two buckets of', and the word *liong* 'two' is a numeral.

9. In Hai-lu Hakka, the suffix -kai represents associative and nominalized relations as the suffix -de in Mandarin The suffix -kai has an allomorph -ai.

(c)*si kai lo su*	'a dead rat'	LM LL LM LM	no sandhi
(d)*fai kai sip ..ıan*	'a bad habit'	LM LL M LL	no sandhi

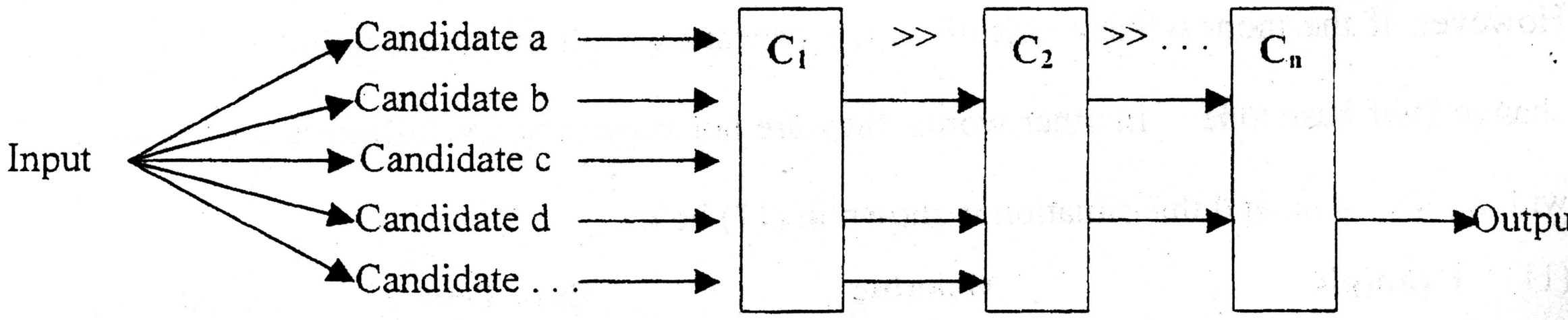
Thus, from the different sandhi phenomena in (10) and (11), it is reasonable to state that th lexicalized type and the -kai type monosyllabic adjectives have two different syntactic identities namely, the former as As, and the later as APs. Furthermore, Hsu (1996) suggests that in Hakk the -ke type[10] construction can be categorized as KP.[11] KP is an functional projection and th suffix -ke is often added after a lexical projection, e.g., NP or AP.

4. CONSTRAINT-BASED APPROACH OF SHANG TONE SANDHI

4.1 Main Concept of Optimality Theory

The Optimality Theory (henceforth, OT) is introduced by Prince and Smolensky in 1993. The main concept of OT is that there are no so-called 'language specific' rules which are extensively used in generative grammar. OT proposes that all grammars of languages are universal and ar constructed with an enormous set of constraints. It is the ranking of the universal constraints tha make languages in the world different from each other, i.e., different languages have differen constraint rankings. Unlike generative grammar, there is no serial derivation process lying in OT framework. Universal grammar provides infinite candidate inputs into the specific languag grammar, and the grammar uses its mechanism called GEN(erator) to evaluate the optimal output with its own constraint ranking. The input-output relation can be diagramed in the schema belov in (12).

(12) OT Schema (Kager, 1999:8)



In the schema above, constraint C_1 is ranked the highest, and it is the least violable constraint. Th

---

10. The suffix -ke in Sixian Hakka spoken in Miaoli County has the same syntactic function as the Hai-lu Hakka suffix -kai .

11. Hsiao (2000 a) uses EP to describe the similar construction in Taiwanese as KP in Hakka.

>> symbol indicates the dominative relation between two constraints. Candidates pass through th constraint evaluator at the same time, but fail to pass the constraints if they can't satisfy th constraints. The candidate, e.g. Candidate d in (12), which satisfies the most constraints wins ou as an optimal output.

4.2 Tonal Constraints for Hai-lu Hakka Shang Tone Sandhi

(13) IDNET-IO-T: the input tones must be identical with the output tones.

IDNET-IO-T is one of the constraints which belongs to the faithfulness constraint family. The faithfulness constraints require that the input elements must be the same as the output elements, and any element change violates this types of constrains.

(14) OCP-T: adjacent identical tones are prohibited.

The OCP-T constraint demands that two identical tones must not be adjacent. It stems from the Obligatory Contour Principle proposed by Leben in 1973, which states that "Adjacent identical tones are banned from the lexical representation of a morpheme."[12]

Tableau (15) and Tableau (16)show how the two constraints interact with each other in Hai-lu Hakka Shang Tone Sandhi.

(15)A. input: 'tiger'

Constraint Ranking: OCP-T >> IDENT-IO-T

lo fu LM LM	OCP-T	IDENT-IO-T
☞ a. LL LM		*
b. LM LM	*!	
c. LL LL	*!	**
☞ d. LM LL		*

B. input: 'ancestors'

Constraint Ranking: OCP-T >> IDENT-IO-T

cu sien LM HM	OCP-T	IDENT-IO-T
☞ a. LM HM		
b. LL HM		*!

(16)A. input: 'tiger'

12. See Kenstowicz (1994:323).

Constraint Ranking: IDENT-IO-T >> OCP-T

lo fu LM LM	IDENT-IO-T	OCP-T
a. LL LM	*!	
☞ b. LM LM		*
c. LL LL	*!*	*
d. LM LL	*!	

B. input: 'ancestors'

Constraint Ranking: IDENT-IO-T >> OCP-T

cu sien LM HM	IDENT-IO-T	OCP-T
☞ a. LM HM		
b. LL HM	*!	

In the tableaux above, the ☞ symbols inside the tableaux indicate the candidates that are evaluated by the constraints as the optimal outputs, while the ☞ symbols outside the tableaux mark the correc output in reality. The * symbol indicates that the candidate violates the constraint, and the symbol after the symbol * shows a fatal violation, which means that the candidate is ruled out i this violation. In (15)A both candidate (a) and candidate (d) are the optimal outputs when OCP-T outranks IDENT-IO-T. In (16)A, however, the optimal output is candidate (b) while the correc tone represented in candidate (a) is ruled out under the ranking of IDENT-IO-T >> OCP-T. Thus the constraint ranking in (15) is more suitable than that in (16). Nevertheless, the two constraint are not enough to account for the Shang Tone Sandhi because there are two optimal outputs i (15)A, and regardless of the ranking in (16)A and (16)B, both rule out the correct outputs. Mor constraints are required to be added in the tableaux.

(17) ALIGN-XP-R: align the right edge of every XP with the right of a tone sandhi domain.

This is a member of the alignment constraint family. It requests the right edge of anXP must align with the right edge of a Tone Sandhi domain. Chen (1996) and Hsiao(2000 a) indicate that Hakka, as well as Min and Mandarin, is a right-prominent language, in other words, the rightmost tone of anXP must be identical with the base tone in a tone sandhi domain. Tableaux (18) and (19 show how the constraint ALIGN-XP-R is ranked in the tableaux.

(18)A. input: 'to hit the tiger'

Constraint Ranking: ALIGN-XP-R >> OCP-T >> IDENT-IO-T

da lo fu]$_{NP}$]$_{VP}$ LM LM LM	ALIGN-XP-R	OCP-T	IDENT-IO-T
☞ a. LL LL LM		*	**
b. LM LM LM	*!*	**	
c. LL LL LL	*!	**	***
d. LM LL LM	*!		*
e. LL LM LM	*!	*	*

In (18)A candidate (c), (d) and (e) violate ALIGN-XP-R once, and candidate (b) violate twice. These candidates thus have fatal violations and are ruled out when they pass the first constraint. Candidate (a) satisfies the ALIGN-XP-R constraint, and can pass onto the next constraint. It is selected as the optimal output though it violates OCP-T once and IDENT-IO-T twice. Violations of lower ranked constraints does not make fatal violations. Tableau (18)A correctly predicts the real sandhi tone for *da lo fu* 'to hit the tiger' as LL LL LM.

B. input: 'to miss the ancestors'

Constraint Ranking: ALIGN-XP-R >> OCP-T >> IDENT-IO-T

siong cu sien]$_{NP}$]$_{VP}$ LM LM HM	ALIGN-XP-R	OCP-T	IDENT-IO-T
a.LL LM HM	*!		*
☞ b. LL LL HM		*	*
c. LM LM HM	*!*	*	
d. LM LL HM	*!		*

In Tableau (18)B, candidate (a) and (d) violate the highest ranked constraint ALIGN-XP-R once, and candidate (c) violates ALIGN-XP-R twice. They are all ruled out at the first violation of ALIGN-XP-R. On the other hand, despite of its violations of OCP-T and IDENT-IO-T, candidate (b) is selected as the optimal output in Tableau (18)B.

(19)A. input: 'to hit the tiger'

Constraint Ranking: OCP-T >> ALIGN-XP-R >> IDENT-IO-T

da lo fu]$_{NP}$]$_{VP}$ LM LM LM	OCP-T	ALIGN-XP-R	IDENT-IO-T
☞ a. LL LL LM	*!		**
b. LM LM LM	*!*	**	
c. LL LL LL	*!*	*	***
☞ d. LM LL LM		*	*
e. LL LM LM	*!	*	*

In Tableau (19)A, the ranking of OCP-T >> ALIGN-XP-R >> IDENT-IO-T rules out candidates (a),

(b), (c), and (e) when they first violate the highest ranked constraint OCP-T, and candidate (d) is selected as the optimal output because it does not violate OCP-T. However, from the inconsistency of the two ☞ symbols inside and outside of the tableau, it is obvious that the constraint ranking in Tableau (19)A is not suitable to account for the sandhi tone for the Shang Tone words.

B. input: 'to miss the ancestors'

Constraint Ranking: OCP-T >> ALIGN-XP-R >> IDENT-IO-T

siong cu sien] NP]VP LM LM HM	OCP-T	ALIGN-XP-R	IDENT-IO-T
☞a.LL LM HM		*	*
☞ b. LL LL HM	*!		*
c. LM LM HM	*!	**	
☞d. LM LL HM		*	*

In Tableau (19)B, candidates (a) and (d) are selected as the optimal outputs, while candidate (b) which represents the actual tone is ruled out because of the fatal violation of OCP-T. The inappropriateness of the ranking OCP-T >> ALIGN-XP-R >> IDENT-IO-T is thus proven again.

5. THE PROBLEM OF SHANG TONE NUMERALS AND CLASSIFIERS

The same as the discussion of Taiwanese classifiers in Hsiao (2000 a), phrases with the Shang Tone classifiers should be considered as CLPs in the Shang Tone Shandhi domain. Furthermore, CLPs are subject to the constraint ALIGN-XP-R. Tableau (20) shows how the sandhi tone for the CLP *sam vong mi ciu* 'three bowls of rice wine' is selected.

(20) input: 'three bowls of rice wine'

Constraint Ranking: ALIGN-XP-R >> OCP-T >> IDENT-IO-T

sam vong]CLP *mi ciu*] NP HM LM LM LM	ALIGN-XP-R	OCP-T	IDENT-IO-T
☞ a. HM LM LL LM			*
b. HM LM LM LM	*!	**	
c. HM LL LL LM	*!	*	**
d. HM LL LL LL	*!*	**	***
e. HM LL LM LM	*!*	*	*

In Tableau (20), candidates (b), (c), (d), and (e) are ruled out at the first violation of the constraint ALIGN-XP-R, and only candidate (a), the optimal output candidate, can pass onto the next constraint. However, the constraints in Tableau (20) are not enough to account for CLPs which contain Shang

Pure numerals, e.g. *liong thung ciang sui* 'two buckets of well water'. The awkward situation i

shown in Tableau (21) below.

(21) input: 'two buckets of well water'

Constraint Ranking: ALIGN-XP-R >> OCP-T >> IDENT-IO-T

liong thung]CLP *ciang sui*]NP LM LM LM LM	ALIGN-XP-R	OCP-T	IDENT-IO-T
☞ a. LL LM LL LM			**
☞ b. LM LM LL LM	*!	*	*
c. LL LL LL LM	*!	***	***
d. LM LM LM LM	*!*	***	
e. LL LM LM LM	*!	**	*

In Tableau (21), candidate (a), which incurs the least violations, is selected as the optimal output, while the actual sandhi tone in candidate (b) violates the highest ranked constraint ALIGN-XP-R once and is thus ruled out. There should be an additional constraint ranked higher than ALIGN-XP-R to allow the Shang Tone numeral words surface with their base tone.

In Hai-lu Hakka, as well as in other Chinese languages, the number character is not stated clearly in normal speech. If the numerals are stated, they will become the focus of the phrases; therefore, their base tone should be preserved. For example, Shang Tone numeral words in the phone number serials as in (22) do not display tone sandhi phenomenon.

(22) Numerals in Phone Number Serials

Example	**Represented Number**	**Base Tone (No Sandhi)**
(a) *kiu ng kiu kiu ng ng kiu*	9599559	LM LM LM LM LM LM LM
(b) *chit ng pat sam it kiu si*	7583194	H LM H HM M LM LL

From the phenomenon above in (22), the constraint in (23) can be added to the constraint ranking in the tableaux.

(23) IDENT-NUMERAL-T: the base tone of numerals in Hai-lu Hakka must be preserved.[13]

Tableau (24) shows how the optimal tone of the phrase *liong thong ciang sui* 'two buckets of well water' is selected.

(24) input: 'two buckets of well water'

13. This constraint is not only suitable for Shang Tone numerals, but also suitable for Yin-ru Tone numerals.

Constraint Ranking: IDENT-NUMERAL-T >> ALIGN-XP-R >> OCP-T >> IDENT-IO-T

liqng thung] CLP *ciang sui*] NP LM LM LM LM	IDENT-NUMERAL -T	ALIGN-XP-R	OCP-T	IDENT-IO-T
☞ a. LM LM LL LM		*	*	*
b. LM LM LM LM		**!	***	
c. LL LL LL LM	*!	**	**	***
d. LL LM LL LM	*!	*		**
e. LL LM LM LM	*!	*	**	*

In Tableau (24), candidates (c), (d) and (e) are ruled out when they first violate the highest ranke constraint IDENT-NUMERAL-T, while candidates (a) and (b) do not incur violations c IDENT-NUMERAL-T, and they pass onto the next constraint ALIGN-XP-R. Candidate (a) and (b both violate the constraint ALIGN-XP-R. Although candidate (a) violates ALIGN-XP-R once, thi do not prevent it to be the optimal output comparing to candidate (b) which violates ALIGN-XP-R twice.

If the phrases composing of Shang Tone numerals and classifiers are regarded as XPs, the sandhi tones can be successfully predicted under the constraint ranking of IDENT-NUMERAL-T > ALIGN-XP-R >> OCP-T >> IDENT-IO-T.

6. CONCLUSION

This paper provides both the derivational rules and the OT constraint ranking to elaborate the Shang Tone Sandhi phenomenon in Hai-lu Hakka. From the discussion above, it is obvious tha the OT framework is more economical to predict the correct tones of Hai-lu Hakka Shang Ton words than the rule-driven steps of Generative discipline. Under the constraint ranking o IDENT-NUMERAL-T >> ALIGN-XP-R >> OCP-T >> IDENT-IO-T, the optimal and actual tones of th Shang Tone words in phrases can be successfully and economically predicted in the tablea transparently in one step.

REFERENCES

Archangeli, Diana, and D. Terence Langendoen, eds. 1997. *Optimality Theory: An Overview.* Oxford: Blackwell.

Chen, Matthew Y. 1996. *Tone Sandhi.* San Diego: University of California.

Gu, Guo-shun, Jian-de Ye, and Shi-song He, eds. 1998. *Xin Zhu Xian Ke Jia Hua Ji Yin Xun Lian.* Hsin-chu: Hsin-chu County Government.

Hsiao, Yuchau E. 2000a. "Optimal Tone Sandhi in Taiwanese." Manuscript. National Chengchi University.

Hsiao, Yuchau E. 2000b. "An Optimality Theory Approach to Hakka Tone Sandhi." Paper fo NACCL 12.

Hsu, Kuei-ping. 1996. "Hakka Tone Sandhi : the Interface between Syntax and Phonology." M.A. thesis, National Tsing Hua University.

Kager, René. 1999. *Optimality Theory.* Cambridge: Cambridge University Press.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar.* Cambridge: Blackwell.

Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese.* Berkeley: University of California Press.

Lin, Huishan. 2000. "Interlanguage Tone Sandhi: Rule interaction in Mandarin-Min Code-Mixing. Paper for NACCL 12.

Luo, Chao-chin. 1984. *Ke Yu Yu Fa.* Taipei: Student.

Luo, Chao-chin. 1990. *Tai Wan De Ke Jia Hua.* Taipei: Taiyuan.

Prince, Alen, and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar.* To appear, MIT Press.

Shih, Chi-ling. 1986. "The Prosodic Domain of Tone Sandhi in Chinese." Ph.D. diss University of California, San Diego.