科技部補助專題研究計畫成果報告
期末報告

從句法音韻介面觀點研究卓蘭饒平三字組及四字組之連讀變調
（第2年）

計 畫 類 別 ： 個別型計畫
計 畫 編 號 ： MOST 103-2410-H-004-086-MY2
執 行 期 間 ： 104年08月01日至105年12月31日
執 行 單 位 ： 國立政治大學語言學研究所

計畫主持人： 蕭宇超

計畫參與人員： 碩士班研究生-兼任助理人員：陳怡臻
碩士班研究生-兼任助理人員：楊雯婷
碩士班研究生-兼任助理人員：洪聖瑋
博士班研究生-兼任助理人員：淩旺楨
博士班研究生-兼任助理人員：黃子權

報 告 附 件 ： 出席國際學術會議心得報告

中 華 民 國 106 年 03 月 30 日

中 文 摘 要 ： 本研究計畫建立了一個大型的卓蘭饒平語料庫，其中包含10662個單音節至四音節樣本，並建置各種音韻標注。在理論上，本研究觀察到，卓蘭饒平的聲調中和由聲調標記、調域OCP及調域信實所啟動。此外，本研究比較「標準韻律理論」與「契合理論」，發現契合制約可能做出錯誤的預測。

中 文 關 鍵 詞 ： 連讀變調、韻律結構、優選理論、卓蘭饒平

英 文 摘 要 ： This research project establishes a large corpus of Zhuolan Raoping, which contains 10662 tokens, monosyllabic to tetrasyllabic, with a variety of phonological coding. Theoretically, this research observes that tone neutralization of Zhuolan Raoping is activated by tone markedness, register OCP and register identity, and a comparison of the standard prosodic theory and the match theory is offered, whereby it shows that the match constraints may render incorrect predictions.

英 文 關 鍵 詞 ： tone sandhi, prosodic structure, OT, Zhuolan Raoping

Trisyllabic and tetrasyllabic tone sandhi in Zhuolan Raoping: a perspective from the syntax-phonology interface

(MOST# 103-2410-H-004-086-MY2)

計畫主持人：蕭宇超

Abstract. This research project establishes a large corpus of Zhuolan Raoping, which contains 10662 tokens, monosyllabic to tetrasyllabic, with a variety of phonological coding. Theoretically, this research observes that tone neutralization of Zhuolan Raoping is activated by tone markedness, register OCP and register identity, and a comparison of the standard prosodic theory and the match theory is offered, whereby it shows that the match constraints may render incorrect predictions.

Keywords: tone sandhi, prosodic structure, OT, Zhuolan Raoping

摘要：本研究計畫建立了一個大型的卓蘭饒平語料庫，其中包含 10662 個單音節至四音節樣本，並建置各種音韻標注。在理論上，本研究觀察到，卓蘭饒平的聲調中和由聲調標記、調域 OCP 及調域信實所啟動。此外，本研究比較「標準韻律理論」與「契合理論」，發現契合制約可能做出錯誤的預測。

關鍵詞：連讀變調、韻律結構、優選理論、卓蘭饒平

1. Introduction

Zhuolan Raoping is a subdialect of Hakka spoken in Miali County, located in central Taiwan. It is surrounded by several bigger Chinese dialects, including Sixian, Hailu, Dongshi, Mandarin, and Taiwanese; it has been greatly affected by those dialects, or more precisely, it is vanishing. This paper offers a corpus analysis of the tone sandhi and tonal phrasing in Zhuolan Raoping. Three questions are pursued in the present research. What are the relevant constraints that interact with each other to achieve the tone sandhi? How are the relative markedness and particular frequencies of the variations predicted through stochastic evaluation? And how are the prosodic theories operated at the syntax-phonology interface?

The rest of this report is organized as follows. §2 discusses the corpus and the tone sandhi. §3 offers an OT analysis of the tone sandhi, where stochastic evaluation, ROE, cophonology and comparative markedness are examined. §4 addresses the corpus and the tonal phrasing, where the standard prosodic theory and the match theory are compared. §5 is the summary.

2. The Corpus

There are six base tones in the tone inventory of this dialect, including four smooth tones and two checked tones. Several studies have investigated the tone values in this dialect, such as Lu (1993), Tu (1998), R. Hsu (2005), Luo (2007), Peng (2007), K. Hsu (2008), M. Cheng and C. Hsu (2013), among others. However, the tone values documented in the literature are not quite consistent, as listed in (1).

(1) Tone inventory of Zhuolan Raoping in previous studies

	Yin Ping	Yang Ping	Shang	Qu	Yin Ru	Yang Ru
Lu (1993)	11	53	31	55	<u>32</u>	<u>5</u>
Tu (1998)	11	53	31	55	<u>2</u>	<u>5</u>
R. Hsu (2005)	11	53	31	55	<u>3</u>	<u>5</u>
Luo (2007)	11	55	53	33	<u>32</u>	<u>55</u>
Peng (2007)	11	53	42	55	<u>22</u>	<u>55</u>
K. Hsu (2008)	11	53	31	55	<u>2</u>	<u>5</u>
M. Cheng and C. Hsu (2013)	31	53	41	43	<u>31</u>	<u>4</u>

In terms of Chao's (1930) tone transcription system, 5 represents high, 3 represents mid, and 1 represents low; 4 can be either high or mid, and 2 can be either mid or low, depending on the tone inventory of a language.

2.1 Tone Inventory

In the present research project, I have redone a series of fieldwork and established a corpus that contains 107 monosyllabic, 1221 disyllabic, 3212 trisyllabic, and 6122 tetrasyllabic tokens. In total, we collect 10662 tokens from 117 hours' recordings, with the help of nine native speakers, five males and four females. The informants are aged between 46 and 73, since the language of the Zhuolan residents that are at or below the age of 40 is often a mixture of Raoping and other Hakka dialects (such as Sixian, Hailu and Dongshi). The recordings are based on designed data, including words, phrases and sentences. According to this fieldwork, I use the notations in (2) to represent the tones in this Hakka dialect.

(2) Tone inventory of Zhuolan Raoping in the present project

	Yin Ping	Yang Ping	Shang	Qu	Yin Ru	Yang Ru
Base Tones	11	53	31	44	$\underline{2}$	$\underline{5}$

There is a mid tone, 33, which occurs only in the surface but does not pertain to the tone inventory. The tone transcriptions in this research project are basically consistent with Tu (1998), R. Hsu (2005), Peng (2007) and K. Hsu (2008). Several studies have looked into the issue of tone structure. Hyman (1993) defines high register as at or above a middle tone height and low register as at or below a middle tone height. Yip (1980, 1989) distinguishes two values of register, [+upper] and [-upper]; the positive value of [upper] includes the range from mid upward, while the negative value from mid downward. Shih (1986) and Inkelas (1987) basically take a similar view of tone height. Yip (1989, 1993, 2002) considers register a mother node that dominates contour nodes, but Duanmu (1990) and Bao (1990, 1999) tone register and contour as sister nodes. In spite of the theory-internal differences, most studies agree that a tone consists of two proportions, register and pitch (or contour). For the purpose of discussion, the numeral tonal notations are specified as in (3); HR stands for high register, and LR for low register; h stands for a high pitch melody, and l for a low pitch melody.

(3) Tones of Zhuolan Raoping

Syllable Types	Smooth				Checked	
Tone Categories	Ping		Shang	Qu	Ru	
	Yin	Yang			Yin	Yang
Base Tones	$11^{[LR, l]}$	$53^{[HR, hl]}$	$31^{[LR, hl]}$	$44^{[HR, h]}$	$\underline{2}^{[LR, l]}$	$\underline{5}^{[HR, h]}$
Sandhi Tones (before HR)	NA	$11^{[LR, l]}$ $33^{[HR, l]}$	$11^{[LR, l]}$	$11^{[LR, l]}$ $33^{[HR, l]}$	$\underline{5}^{[HR, h]}$	$\underline{2}^{[LR, l]}$
Sandhi Tones (before LR)	NA	$33^{[HR, l]}$	$33^{[HR, l]}$	$33^{[HR, l]}$	$\underline{5}^{[HR, h]}$	NA

A smooth tone, or a long tone, is a tone borne by an open syllable or a syllable that ends in a sonorant. A checked tone, or a short tone, is a tone borne by a closed syllable which ends in a voiceless stop. The status of checked tones in Chinese dialects has been controversial. Some studies (Zee & Maddieson 1979, Xu et al. 1981, Qian 1992, Li 2003) suggest that there is an allophonic connection between checked tones and smooth tones, while others (Ping 2001, Huang 2011) argue for an asymmetry between the two types of tones. To avoid digression, this report focusses on tone sandhi of the smooth tones. In (3), I posit two high-register smooth tones, 53 and 44, and two low-register ones, 31 and 11. I consider the surface mid tone, 33, as the low end of the high register, i.e., [HR, l].

2.2 Tone Sandhi

Two patterns of the smooth tone sandhi can then be generalized. The first pattern observed is that the low-register 31 maps to the high-register 33 before a low-register tone, as in (4), but to the low-register 11 before a high-register tone, as in (5).

(4) 31-LR → 33-LR

a.	*coi*	*do*	‘kitchen knife’	b.	*fo*	*cha*	‘train’	
	31	11			31	11		
	[LR, hl]	[Lr, 1]			[LR, hl]	[LR, 1]		input
	33	11			33	11		
	[HR, l]	[Lr, 1]			[HR, l]	[LR, 1]		output

c.	*sia*	*ho*	‘written’	d.	*se*	*zii*	‘child’	
	31	31			31	31		
	[LR, hl]	[Lr, h1]			[LR, hl]	[LR, h1]		input
	33	31			33	31		
	[HR, l]	[Lr, h1]			[HR, l]	[LR, h1]		output

e.	*cai*	*sed*	‘colorful’	f.	*gi*	*zed*	‘season’	
	31	$\underline{2}$			31	$\underline{2}$		
	[LR, hl]	[Lr, 1]			[LR, hl]	[LR, 1]		input
	33	$\underline{2}$			33	$\underline{2}$		
	[HR, l]	[Lr, 1]			[HR, l]	[LR, 1]		output

(5) 31-HR → 11-HR

a.	*go*	*ngien*	‘New Year’	b.	*giam*	*pui*	‘to lose weight’	
	31	53			31	53		
	[LR, hl]	[HR, hl]			[LR, hl]	[HR, hl]		input
	11	53			11	53		
	[LR, l]	[HR, hl]			[LR, l]	[HR, hl]		output

c.	*gong*	*fa*	‘to talk’	d.	*den*	*lu*	‘gift’	
	31	44			31	44		
	[LR, hl]	[HR, h]			[LR, hl]	[HR, h]		input
	11	44			11	44		
	[LR, l]	[HR, h]			[LR, l]	[HR, h]		output

e. *co* *iog* ‘medicine’ f. *pui* *hab* ‘to cooperate’

31 5		31 5		
[LR, hl] [HR, hl]		[LR, hl] [HR, h]		input
11 5		11 5		
[LR, l] [HR, hl]		[LR, l] [HR, h]		output

In the corpus, there are 458 [31 LR] combinations in disyllabic, trisyllabic, and tetrasyllabic tokens, as in (6); in 444 of them 31 maps to 33, found in 96.94%, while only in 14 of them 31 maps to 11, found in 3.06%. On the other hand, there are 1068 [31 HR] combinations in disyllabic, trisyllabic, and tetrasyllabic tokens, as in (7); in only 71 of them 31 maps to 33, found in 0.66%, but in 997 of them 31 maps to 11, found in 99.34%. These numbers indicate that the sandhi of the low-register 31 conforms to the OCP effects on the register tier quite consistently.

(6) 31 LR (X = any tone)

	31 → 33	31 → 11	Total
[31 LR]	98	0	98
[31 LR X]	50	2	52
[X 31 LR]	48	1	49
[31 LR X X]	89	6	95
[X 31 LR X]	80	3	83
[X X 31 LR]	79	2	81
Total	444	14	458
Percentage	96.94%	3.06%	100%

(7) 31 HR (X = any tone)

	31 → 33	31 → 11	Total
[31 HR]	2	220	222
[31 HR X]	9	90	99
[X 31 HR]	4	73	77
[31 HR X X]	12	286	298
[X 31 HR X]	25	147	172
[X X 31 HR]	19	181	200
Total	71	997	1068
Percentage	0.66%	99.34%	100%

The second pattern of the smooth tone sandhi observed in this paper is that the high-register 53 and 44 may map to 33 before a low-register tone, as in (8-9), or map to the 11 or 33 before a high-register tone, as in (10-11).

(8) 53-LR → 33-LR

a.	*ngin*	*sen*	'life'	b.	*pen*	*iu*	'friend'	
	53	11			53	11		
	[HR, hl]	[LR, 1]			[HR, hl]	[LR, 1]		input
	33	11			33	11		
	[HR, l]	[LR, 1]			[HR, l]	[LR, 1]		output
c.	*teu*	*den*	'first class'	d.	*fin*	*hong*	'beside'	
	53	31			53	31		
	[HR, hl]	[LR, h1]			[HR, h]	[LR, h1]		input
	33	31			33	31		
	[HR, l]	[LR, h1]			[HR, l]	[LR, h1]		output
e.	*fung*	*sed*	'red'	f.	*shi*	*zed*	'time'	
	53	$\underline{2}$			53	$\underline{2}$		
	[LR, hl]	[Lr, 1]			[LR, hl]	[LR, 1]		input
	33	$\underline{2}$			33	$\underline{2}$		
	[HR, l]	[Lr, 1]			[HR, l]	[LR, 1]		output

(9) 44-LR → 33-LR

a.	*shu*	*gin*	'root'	b.	*nguai*	*sam*	'coat'	
	44	11			44	11		
	[HR, h]	[LR, 1]			[HR, h]	[LR, 1]		input
	33	11			33	11		
	[HR, l]	[LR, 1]			[HR, l]	[LR, 1]		output
c.	*ten*	*iang*	'movie'	d.	*pu*	*zhong*	'minister'	
	44	31			44	31		
	[HR, h]	[LR, h1]			[HR, h]	[LR, 1]		input
	33	31			33	31		
	[HR, l]	[LR, h1]			[HR, l]	[LR, 1]		output
e.	*tai*	*sed*	'snow'	f.	*ti*	*id*	'first'	
	44	$\underline{2}$			44	$\underline{2}$		
	[HR, h]	[LR, 1]			[HR, h]	[LR, 1]		input
	33	$\underline{2}$			33	$\underline{2}$		
	[HR, l]	[LR, 1]			[HR, l]	[LR, 1]		output

(10) 53-HR → 33/11-HR

a.	*mo*	*cen*	‘no money’	b.	*cen*	*tu*	‘future’	
	53	53			53	53		
	[HR, hl]	[HR, h1]			[HR, hl]	[HR, h1]		input
	33	53			33	53		
	[HR, l]	[HR, h1]			[HR, l]	[HR, h1]		output 1
	11	53			11	53		
	[LR, l]	[HR, h1]			[LR, l]	[HR, h1]		output 2

c.	*fung*	*teu*	‘red beans’	d.	*vo*	*shong*	‘monk’	
	53	44			53	44		
	[HR, hl]	[HR, h]			[HR, hl]	[HR, h]		Input
	33	44			33	44		
	[HR, l]	[HR, h]			[HR, l]	[HR, h]		output 1
	11	44			11	44		
	[LR, l]	[HR, h]			[LR, l]	[HR, h]		output 2

e.	*tong*	*hog*	‘classmate’	f.	*liu*	*hog*	‘to study abroad’	
	53	5			53	5		
	[HR, hl]	[Hr, h]			[HR, hl]	[Hr, h]		input
	33	5			33	5		
	[HR, l]	[Hr, h]			[HR, l]	[Hr, h]		output 1
	11	5			11	5		
	[LR, l]	[HR, h]			[LR, l]	[HR, h]		output 2

(11) 44-HR → 33/11-HR

a.	*heu*	*mun*	‘rear door’	b.	*van*	*cen*	‘change money’	
	44	53			44	53		
	[HR, h]	[HR, h1]			[HR, h]	[HR, h1]		input
	33	53			33	53		
	[HR, l]	[HR, h1]			[HR, l]	[HR, h1]		output 1
	11	53			11	53		
	[LR, l]	[HR, h1]			[LR, l]	[HR, h1]		output 2

c.	*hen* *toi*	‘modern’	d.	*kiung* *ha*	‘together’	
	44 44			44 44		
	[HR, h] [Hr, h]			[HR, h] [Hr, h]		input
	33 44			33 44		
	[HR, l] [Hr, h]			[HR, l] [Hr, h]		output 1
	11 44			11 44		
	[LR, l] [HR, h]			[LR, l] [HR, h]		output 2

e.	*tai* *lid*	‘big strength’	f.	*kiu* *led*	‘lunar calendar’	
	44 <u>5</u>			44 <u>5</u>		
	[HR, h] [Hr, h]			[HR, h] [Hr, h]		input
	33 <u>5</u>			33 <u>5</u>		
	[HR, l] [Hr, h]			[HR, l] [Hr, h]		output 1
	11 <u>5</u>			11 <u>5</u>		
	[LR, l] [HR, h]			[LR, l] [HR, h]		output 2

The corpus contains 502 [53 LR] combinations in disyllabic, trisyllabic, and tetrasyllabic tokens, as in (12); in 498 of them 53 maps to 33, found in 99.20%, while only in 4 of them 53 maps to 11, found in 0.80%. There are 643 [44 LR] combinations in disyllabic, trisyllabic, and tetrasyllabic tokens, as in (13); in 631 of them 44 maps to 33, found in 98.13%, while only in 12 of them 44 maps to 11, found in 1.87%.

(12) 53 LR (X = any tone)

	53 → 33	53 → 11	Total
[53 LR]	78	0	78
[53 LR X]	49	1	50
[X 53 LR]	66	0	66
[53 LR X X]	104	2	106
[X 53 LR X]	117	1	118
[X X 53 LR]	84	0	84
Total	498	4	502
Percentage	99.20%	0.80%	100%

(13) 44 LR (X = any tone)

	44 → 33	44 → 11	Total
[44 LR]	82	1	83
[44 LR X]	48	3	51
[X 44 LR]	43	1	44

[44 LR X X]	154	4	158
[X 44 LR X]	160	2	162
[X X 44 LR]	144	1	145
Total	631	12	643
Percentage	98.13%	1.87%	100%

On the other hand, there are 836 [53 HR] combinations in disyllabic, trisyllabic, and tetrasyllabic tokens, as in (14); in 652 of them 53 maps to 33, found in 77.99%, while in 184 of them 53 maps to 11, found in 22.01%. There are 992 [44 HR] combinations in disyllabic, trisyllabic, and tetrasyllabic tokens, as in (15); in 729 of them 44 maps to 33, found in 73.49%, while in 263 of them 44 maps to 11, found in 26.51%.

(14) 53 HR (X = any tone)

	53 → 33	53 → 11	Total
[53 HR]	127	19	146
[53 HR X]	54	27	81
[X 53 HR]	51	15	66
[53 HR X X]	119	61	180
[X 53 HR X]	138	40	178
[X X 53 HR]	163	22	185
Total	652	184	836
Percentage	77.99%	22.01%	100%

(15) 44 HR (X = any tone)

	44 → 33	44 → 11	Total
[44 HR]	96	32	128
[44 HR X]	60	20	80
[X 44 HR]	47	18	65
[44 HR X X]	187	81	268
[X 44 HR X]	147	59	206
[X X 44 HR]	192	53	245
Total	729	263	992
Percentage	73.49%	26.51%	100%

Accordingly, 53 and 44 tend to map to 33 even before a high-register tone, found in over 75.55% of the data. In spite of that, there are around 24.45% of the data where 53 and 44 map to 11 before high-register tones. The average percentages of the high-register tone variations are calculated as in (16).

(16) 53/44 HR

	→ 33	→ 11	Total
…53 HR…	652	184	836
…44 HR…	729	263	992
Total	1381	447	1828
Percentage	≈75.55%	≈24.45%	100%

The smooth low tone, 11, in this dialect does not undergo tone sandhi, neither before a low-register tone nor before a high-register tone, as in (17-18).

(17) 11-LR → 11-LR

a.	*dung*	*fung*	‘East wind’	b.	*san*	*hang*	‘pits’	
	11	11			11	11		
	[LR, l]	[Lr, 1]			[LR, l]	[Lr, 1]		input
	11	11			11	11		
	[LR, l]	[Lr, 1]			[LR, l]	[Lr, 1]		output
c.	*shin*	*ti*	‘body’	d.	*ten*	*shiu*	‘help’	
	11	31			11	31		
	[LR, l]	[Lr, h1]			[LR, l]	[Lr, h1]		input
	11	31			11	31		
	[LR, l]	[Lr, h1]			[LR, l]	[Lr, h1]		output
e.	*ciang*	*sed*	‘green’	f.	*ge*	*ngiug*	‘chicken meat’	
	11	<u>2</u>			11	<u>2</u>		
	[LR, l]	[LR, 1]			[LR, l]	[LR, 1]		input
	11	<u>2</u>			11	<u>2</u>		
	[LR, l]	[LR, 1]			[LR, l]	[LR, 1]		output

(18) 11-HR → 11-HR

a.	*sim*	*cin*	‘mood’	b.	*kiu*	*gung*	‘uncle’	
	11	53			11	53		
	[LR, l]	[HR, hl]			[LR, l]	[HR, hl]		input
	11	53			11	53		
	[LR, l]	[HR, hl]			[LR, l]	[HR, hl]		output

c.	*ten*	*fa*	‘to listen’	d.	*mi*	*mi*	‘deal’		
	11	44			11	44			
	[LR, l]	[HR, h]			[LR, l]	[HR, h]			input
	11	44			11	44			
	[LR, l]	[HR, h]			[LR, l]	[HR, h]			output
e.	*sen*	*fad*	‘life’	f.	*man*	*ngied*	‘full moon’		
	11	5			11	5			
	[LR, l]	[HR, h]			[LR, l]	[HR, h]			input
	11	5			11	5			
	[LR, l]	[HR, h]			[LR, l]	[HR, h]			output

The corpus also shows that given a string of tones, the rightmost tone is preserved, as in (4-18). Such preservation of tone is conditioned within the domain of phonological phrase. The prosodically conditioned tone sandhi will be discussed in section 4.

3. An OT Analysis of Tone Sandhi

According to the corpus, we can summarize four phenomena of the smooth tone sandhi. First, 31, 53 and 53 are neutralized as either 33 or 11 in non-final position. Second, 11 does not undergo tone sandhi. Third, any tone in the rightmost remains unchanged. Finally, 11 is derived before a high-register tone, whereas 33 is derived either before a low-register tone or before a high-register tone.

3.1 Tonal Constraints

These phenomena can be achieved through interactions between constraints of tone markedness and faithfulness, in terms of *Optimality Theory* (hereafter OT, Prince and Smolensky 1993/2004). Three sets of constraints are proposed below.

(19) Tone Markedness (TMARKS)

a. *31: assign a violation mark for every low-falling tone, 31, in the output.

b. *53: assign a violation mark for every high-falling tone, 53, in the output.

c. *44: assign a violation mark for every high tone, 44, in the output.

d. *11: assign a violation mark for every low tone, 11, in the output.

(20) Tone Faithfulness

a. IDENT-T (ID-T): assign a violation mark for every tone in the output that is not identical to its input.

b. IDENT-$T_{]\varphi}$ (ID-$T_{]\varphi}$): assign a violation mark for every φ-final tone in the output that is not identical to its input.

(21) OCP Effects

a. *LR-LR: assign a violation mark for every pair of adjacent low-register tones in the output.

b. *HR-HR: assign a violation mark for every pair of adjacent high-register tones in the output.

The obligatory sandhi of 31, 53 and 44 suggests that the relevant tone markedness constraints, as in (19a-c), are ranked higher than ID-T, as in (20a). The preservation the rightmost tone indicates that ID-$T_{]\varphi}$, as in (20b), dominates the tone markedness constraints in (19a-c). The symbol, φ, refers to the domain of the phonological phrase, which will be discussed in section 4. The following tableaux show how these constraints work.

(22) Input: 31 31 Output: 33 31 = (4c,d)

	ID-$T_{]\varphi}$	*31	ID-T
☞a. 33 31			*
b. 31 31		*!*	
c. 31 33	*!		*

(23) Input: 53 31 Output: 33 31 = (8c,d)

	ID-$T_{]\varphi}$	*31, *53	ID-T
☞a. 33 31			*
b. 53 31		*!*	
c. 53 33	*!		*

(24) Input: 44 31 Output: 33 31 = (9c,d)

	ID-$T_{]\varphi}$	*31, *44	ID-T
☞a. 33 31			*
b. 44 31		*!*	
c. 44 33	*!		*

The absence of 31, 53 and 44 in non-final position indicates that they are more marked than 11 in this language. The lack of 11 sandhi suggests that *11 is bottom-ranked in this language, as in (25).

(25) Input: 11 31   Output: 11 31   = (10a,b)

	ID-T$_{]\varphi}$	*31	ID-T	*11
a. 33 31		*	*!	
☞b. 11 31		*		*
c. 11 33	*!		*	*

The constraints in (19) and (20) obligate 31, 53 and 44 to change in non-final position. Whether they map to 33 or 11 is subsumed under the OCP effects. When OCP operates on the low-register tier, no surface low-register tone is allowed to occur before any low-register tone. The constraint *LR-LR requires 31, 53 and 44 to map to 33 but not to 11 before a low-register tone.

(26) Input: 31 31   Output: 33 31   = (4c,d)

	ID-T$_{]\varphi}$	*LR-LR	*31	ID-T
☞a. 33 31				*
b. 11 31		*!	*	*
c. 31 31		*!	**	
d. 31 33	*!			*

(27) Input: 53 31   Output: 33 31   = (8c,d)

	ID-T$_{]\varphi}$	*LR-LR	*31, *53	ID-T
☞a. 33 31				*
b. 11 31		*!	*	*
c. 53 31			*!*	
d. 53 33	*!		*	*

(28) Input: 44 31   Output: 33 31   = (9c,d)

	ID-T$_{]\varphi}$	*LR-LR	*31, *44	ID-T
☞a. 33 31			*	*
b. 11 31		*!	*	
c. 44 31			**!	*
d. 44 33	*!		*	*

On the other hand, when OCP operates on the high-register tier, no surface high-register tone is allowed to occur before any high-register tone. The constraint *HR-HR requires 31, 53 and 44 to map to 11 instead of 33 before a high-register tone. I will shows later the ranking of *HR-HR is much lower than TMARKS that The corpus shows that 31 observes *HR-HR quite consistently, as in (29).

(29) Input: 31 53   Output: 11 53   = (5a,b)

	ID-T]φ	*31,*53	*HR-HR	ID-T
a. 33 53		*	*!	*
☞b. 11 53		*		*
c. 31 53		**!		
d. 31 33	*!			*

However, the sandhi of 53 and 44 displays variation. As discussed in section 2, the corpus shows that in about 75.55% of the data 53 and 44 map to 33, while in about 24.45% of the data they map to 11. Accordingly, register faithfulness also plays a role.

(30) Register Faithfulness

a. IDENT-HIGH-REGISTER (ID-HR): Assign a violation mark for every high register in the output that is not identical to its input.

a. IDENT-LOW-REGISTER (ID-LR): Assign a violation mark for every low register in the output that is not identical to its input.

The high-register tone mappings are determined by the interaction between ID-HR and *HR-HR. The percentages in (16) can then be interpreted as (31).

(31) Flexible constraint rankings

75.55% 33-mapping: ID-HR >> *HR-HR

24.45% 11-mapping: *HR-HR >> ID-HR

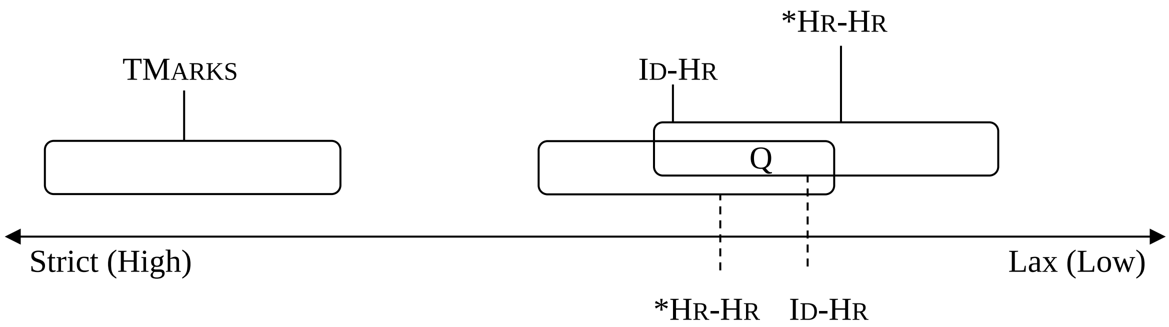
In other words, OCP-HR is 78.01% probably outranked by ID-HR, but 21.99% probably outranks the latter. As to ID-LR, it is bottom-ranked in this dialect, since 31, 53 and 44 consistently map to 33 before a low-register tone.

In classical OT, Eval chooses an optimal output, and eliminates all the remaining candidates without differentiating the well-fornedness degrees among them. The device that handles language variation is constraint reranking (Kiparsky, 1993; Kroch, 1994). Subsequent research has attempted to describe the conditions and frequencies in language variation. In the following sections, I will examine three recent approaches to the 33/11 variations, including *Stochastic OT*, *Rank-Ordering Model of Eval*, and *Cophonology*.

3.2 Stochastic OT

Stochastic evaluation (Boersma and Hayes 2001) is based on an unbounded, continuous scale of constraint strictness, where higher values correspond to higher-ranked, stricter constraints. Each constraint is arbitrarily assigned a selected point (i.e., the value used at evaluation time) on the ranking scale. As illustrated in (10), the selected point is not a single point, but is associated with a range of values, represented by the rounded rectangles. The center of the range is referred to as the ranking value. If the ranges of the selected points do not overlap, the ranking between the relevant constraints is categorical; at this point, TMARKS (*31 *53 and *44) constantly outrank ID-HR. If the ranges overlap, the constraint ranking is variable; Q indicates the overlapped area between ID-HR and *HR-HR, in which ID-HR may choose a part that is lower than *HR-HR, and a reverse ranking may occur, as indicated by the vertical dashed line.

(32) Ranking scale



The constraint ranges are interpreted as probability distributions and described with a normal (= Gaussian) distribution, which has a peak in the center and drops toward zero on both sides. A normal distribution is expressed by its mean (μ). 68.27% of the values drawn from a normal distribution lie within one standard deviation (σ) from the mean (i.e., between μ - σ and $\mu + \sigma$), 95.45% within two σ's, and 99.73% with three σ's (≈ 100%). The value for one standard deviation is 2.0. The calculation formula set up in the computer for a constraint range is as follows (Lu 2013, 2016).

(33) A constraint range = 12

	2	(Standard deviation)
x	3σ's	(≈ 100%)
x	2	(Both sides of the peak)
	12	

After 10662 data, the computer assigns the ranking value of TMARK as 121.56, and that of ID-HR as 92.89 such that the distance between these two kinds of constraints prevents their ranges from overlapping. Consider again (31), which indicates that ID-HR is 24.45% probably dominated by *HR-HR. Accordingly, 48.90% (two times of 24.45%) of the two constraint ranges are overlapped. The ranking value for *HR-HR is thus calculated as 86.89. Details are given in (34).

(34) Relevant values (TMARKS = *31, *53, *44)

Ranking value of TMARKS:	121.56
Ranking value of ID-HR:	92.89
Bottom value of ID-HR:	86.89 (92.89 – 6)
Overlapped area:	5.87 (12*48.90%)
Top value of *HR-HR:	92.76 (86.89 + 5.87)
Ranking value of *HR-HR:	86.76 (92.76 – 6)

If the value of ID-HR randomly falls on 92.77 (or any point higher than the top value of *HR-HR) and the value of *HR-HR falls on 86.88 (or any point lower than the bottom value of ID-HR), ID-HR will dominate *HR-HR, a ranking that results in the common 33-mapping. The probability of this ranking is 75.55% in the corpus. On the other hand, if the value of *HR-HR randomly falls on 92.76 (or any point near its top value) and the value of ID-HR falls on 86.89 (or any point near its bottom value), *HR-HR will dominate ID-HR, a ranking that obtains the less frequent 11-mapping. The probability of this ranking is 24.45% in the corpus. In brief, allowing variable rankings between constraints, stochastic evaluation makes predictions of both particular frequencies and ranking adjustments.

(35) a. 75.55%: ID-HR (92.77 or higher) >> *HR-HR (86.88 or lower)

Input: 44[HR, h] 53[HR, hl] Output: 33[HR, l] 53[HR, hl] (common reading)

	*44, *53	ID-HR	*HR-HR
a.44[HR, h] 53[HR, hl]	**!		*
☞b.33[HR, l] 53[HR, hl]	*		*
c.11[LR, l] 53[HR, hl]	*	*!	

b. 28.38%: *HR-HR (92.76 or near) >> ID-HR (86.89 or near)

Input: 44[HR, h] 53[HR, hl] Output: 11[LR, l] 53[HR, hl] (less frequent reading)

	*44, *53	*HR-HR	ID-HR
a.44[HR, h] 53[HR, hl]	**!	*	
b.33[HR, l] 53[HR, hl]	*	*!	
☞c.11[LR, l] 53[HR, hl]	*		*

The grammar hierarchy can be sketched as (36), where the shaded boxes indicate that the ranges of ID-HR and *HR-HR have an overlap. The ranking values are calculated and listed in the right column.

(36) Grammar hierarchy

	Ranking values
*ML,*H, *HM	121.56
|	
IDENT-HR	92.89
*HR-HR	86.76

The basic purpose of stochastic evaluation is to operate on the Gradual Learning Algorithm (GLA) developed by Boersma (1997, 1998). The algorithm compares the learning datum and the grammar-generated form; when there is a mismatch, constraint demotion (Tesar and Smolensky 1998, 2000) occurs to obtain the learning form. The high-register tone mappings in Zhuolan Raoping can be captured in a similar fashion. As mentioned earlier, Zhuolan Raoping is closely surrounded by several major Chinese dialects. Previous studies (J. Hsu 2005, Peng 2007, and K. Hsu 2008) observed an exclusive rule that the high-register tones map to 11 before any high-register tone, but I have found in the present corpus that the mappings of the high-register tones display variation, and they tend to map to 33. I suggest that there are at least three possible factors for the emergence of 33. First, it may be attributed to the influences of the neighboring dialects. Some relevant tone sandhi patterns in those dialects are observed below.

(37) 33-mapping in surrounding dialects

a. Taiwanese: 55 → 33
13 → 33

b. Dongshi: 35 → 33

Due to the long-term, close interaction, the common 33-mappings in the neighboring dialects are likely to affect Zhuolan Raoping. Second, there are many syllables in Zhuolan Raoping that incorporate readings from other dialects. Examples are given in (38).

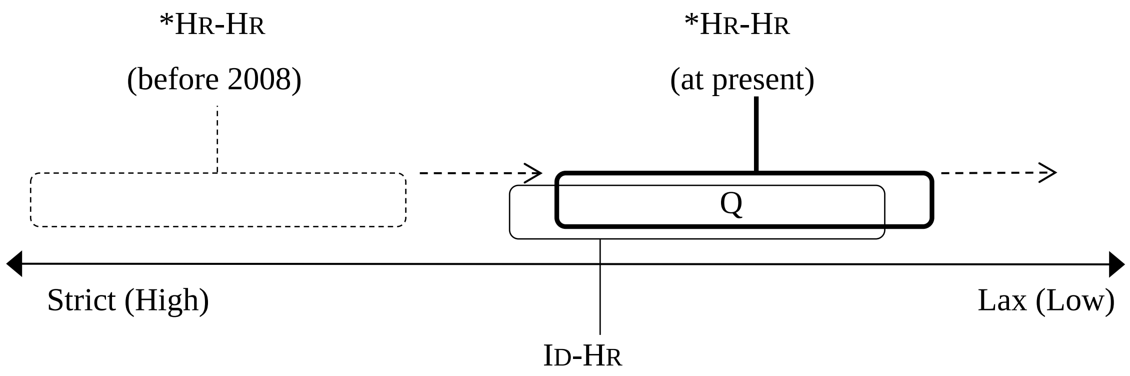
(38) Words with alternative readings

	Native	Language contact
moi ‘sister’	31	55 (from Sixian)
gung ‘old man’	11	53 (from Hailu)
kiag ‘drama’	3	5 (from Taiwanese)

A low-register tone in Zhuolan Raoping consistently maps to a low tone before a high-register tone, as in (1b). Interestingly, many words in this dialect display alternative readings due to language contact, and some low-register tone words develop high-register tone variants, as in (38), where readings from other dialects are adopted alternatively. In that event, the high-register tone variants are subject to the interaction between *HR-HR and ID-HR as well, and the 33-mapping increases.

Finally, this dialect may display a case of tonal analogy, as proposed by Huang and Hsiao (2016). The pre-HR 11-mapping is later reanalyzed as 33-mapping, which is extended to the sandhi position before all contexts. I am not intended to pinpoint the influence of any specific factor on the sandhi variation, but to indicate that the sandhi of the high-register tones is in transition, particularly, from 11-mapping to 33-mapping. In other words, the selected point of the constraint, *HR-HR, has gradually moved downward on the continuous scale ranking, as schematized in (39)

(39) Constraint demotion



The dashed rounded rectangle represents the selected point of *HR-HR before 2008, where the categorical ranking "*HR-HR >> ID-HR" renders the exclusive 11-mapping observed in the previous studies. The boldfaced rounded rectangle represents the selected point of *HR-HR at present; the overlap between ID-HR and *HR-HR allows the acquired 33-mapping to emerge, as found in the current corpus. The right-headed dashed arrows indicate the direction of the constraint movement. The schema also implies that *HR-HR will be continuously demoted and the 11-mapping may eventually vanish.

3.3 Rank-Ordering Model of Eval

The *Rank-Ordering Model of Eval* (ROE) is proposed by Coetzee (2006). This model requires all constraints to evaluate the full candidate set and allows Eval to impose a well-formedness rank-ordering on the full candidate set. The candidate that is on the higher position of the rank-ordering is the more unmarked, and is more likely to be selected as output. No constraint reranking is assumed. The constraint set is divided into two strata by a line of ‘critical cut-off’. The constraints that are ranked above the cut-off line behave as the same as those in classical OT; violations of these constraints are fatal, and only one optimal output is chosen when the rest of the candidates are disfavored. Conversely, the constraints that are ranked below the cut-off line do not rule out candidates, but simply force a harmonic rank-ordering on the candidates that survive below the cut-off line. Violations of these constraints are not fatal, but they determine the relative markedness of the candidates that emerge. When there is more than one candidate that infringes only the constraints from below the cut-off line, variant outputs will be selected. As Coetzee indicates, there are three logical possibilities regarding the critical cut-off. First, there are multiple candidates that violate the constraints from below the cut-off line, and surface variants occur. Second, only one candidate that is only disfavored by the constraints from below the cut-off line, and it emerges as the sole output. Third, all candidates violate at least one constraint that ranks above the cut-off line, and only the best candidate is selected as the output but no variant is recognized.

The sandhi of the high-register tones in Zhuolan Raoping can be under the influence of the critical cut-off line.

(40) Input: $44^{[\text{HR, h}]}$ $53^{[\text{HR, hl}]}$ Output: $33^{[\text{HR, l}]}$/$11^{[\text{LR, l}]}$ $53^{[\text{HR, hl}]}$

	*44, *53	ID-HR	*HR-HR
a. $44^{[\text{HR, h}]}$ $53^{[\text{HR, hl}]}$	**!		*
☞b. $33^{[\text{HR, l}]}$ $53^{[\text{HR, hl}]}$	*		*
☞c. $11^{[\text{LR, l}]}$ $53^{[\text{HR, hl}]}$	*	*	

In (40), candidate (a) is ruled out by TMARK. Candidates (b) and (c) incur a violation of either ID-HR or *HR-HR; since both constraints are ranked below the critical cut-off line, both candidates emerge. ROE is able to indicate that candidate (b) is more unmarked than candidate (c), as the violation of the higher-ranked ID-HR is more serious than that of the lower-ranked *HR-HR. However, ROE is unable to make predictions of particular frequencies; the 75.55% M-mappings and the 24.45% L-mappings of the high-register tones in the corpora cannot be captured. Another disadvantage of ROE is that there is no prediction of the emergence of the acquired form (such as the acquired M-mapping), unlike stochastic evaluation, which assumes constraint ranking adjustments.

3.4 Cophonology

The theory of *Cophonology* is developed in work by Orgun (1996), Anttila (1997), Inkelas and Zoll (2007), and others. This theory permits co-existing distinct phonological systems in a language, known as cophonologies. Cophonologies of a language are related in a grammar hierarchy whose superordinate node is referred to as the “Master Ranking”, as depicted in (41).

(41)

Master Ranking
Con X, Con Y

Cophonology A — Con X >> Con Y

Cophonology B — Con Y >> Con X

The Master Ranking includes a partial ranking of constraints which all individual cophonologies in the language must conform to. Cophonologies have the same set of phonological constraints, but may be related to different constraint rankings. In this sense, we could posit two cophonologies for the high-register tone mappings in Zhuolan Raoping, as in (42).

(42) Cophonologies

Cophonology A: Id-Hr >> *Hr-Hr

Cophonology B: *Hr-Hr >> Id-Hr

Inkelas and Zoll (2007) have indicated that cophonologies are indexed to individual morphological constructions, each of which can be associated with its own constraint ranking. Three problems are then in order. First, the 11/33 variations in Zhuolan Raoping are not structure-related; both readings may occur in the same structure, or in the same word.[1] Second, the cophonology analysis offers no information of relative markedness or prediction of particular frequencies. Finally, it does not predict the emergence of the learning form either.

3.5 Comparative Markedness

A problem arises. The ranking of *11 below Id-T is expected to predict the "11 → 11" mapping, but this expectation is not met.

(43) Input: $11^{[\text{Lr}, \text{l}]}\ 11^{[\text{Lr}, \text{l}]}$ Output: $11^{[\text{Lr}, \text{l}]}\ 11^{[\text{Lr}, \text{l}]}$ = (17a,b)

	*Lr-Lr	*31	Id-T	Id-Lr	*11
(☞)a. $11^{[\text{Lr}, \text{l}]}\ 11^{[\text{Lr}, \text{l}]}$	*!				**
*☞b. $31^{[\text{Hr}, \text{l}]}\ 11^{[\text{Lr}, \text{l}]}$		*	*	*	*

In (43), candidate (b), in which the left 11 maps to 33, is incorrectly selected, as

[1] Tone sandhi in languages like Tiantou is sensitive to morphosyntactic structures, as in (30a,b), which may possibly be accounted for by the theory of Cophonology. In particular, the VO and non-VO structure can be related to separate cophonologies (Hsiao 2015b).

indicated by the asterisked hand syllable, *☞; whereas the actual optimal output, candidate (b), in which both 11's remain unchanged, is inevitably eliminated by *LR-LR, as indicated by the parenthesized hand syllable, (☞). In this section, I will suggest that the problem can be solved by imposing the grandfathering effects on OCP.

McCarthy (2003) suggests that universal constraints may specify the *locus of violation* in addition to the number of violation marks. He indicates that every markedness constraint, M_i, "is defined by some locus function LOC_i. The result of applying M_i to *cand* is a number of violation-marks equal to the cardinality of the set obtained by applying LOC_i to *cand* (2003: 7)." LOC_i are not strings, but individual tones. Hsiao (2015b) includes the specification of violation locus in the definition of OCP and distinguishes between OCP-$\alpha^{[LOC1]}$ and OCP-$\alpha^{[LOC2]}$. Given a bi-tonal sequence, α_j α_k, OCP-$\alpha^{[LOC1]}$ is violated by α_j but not α_k, while OCP-$\alpha^{[LOC2]}$ is violated by α_k but not α_j.

In the case of Zhuolan Raoping, the constraint *LR-LR should be specified as *LR-LR$^{[LOC1]}$; namely, the locus of violation is the low register of the left tone. In reaction to the problem of (43), I would like to propose here that the grandfathering effect is related to the locus of violation. In particular, it is necessary to posit both old and new *LR-LR$^{[LOC1]}$ constraints, namely, *LR$_{OLD}$-LR$^{[LOC1]}$ and *LR$_{NEW}$-LR$^{[LOC1]}$. *LR$_{OLD}$-LR$^{[LOC1]}$ is violated by a LR$_{OLD}$-LR pair, where the low register of the left tone is old; * LR$_{NEW}$-LR$^{[LOC1]}$ is violated by a LR$_{NEW}$-LR pair , where the low register of the left tone is new. In this sense, whether the low register of the right tone is old or new does not matter. The "*LR$_{NEW}$-LR$^{[LOC1]}$ >> ID-LR >> *LR$_{OLD}$-LR$^{[LOC1]}$" ranking enables a [11$_{OLD}$ 11] sequence to surface and efficiently excludes a derived [11$_{NEW}$ 11] sequence, as illustrated below.

(44) Input: 11$^{[LR, l]}$ 11$^{[LR, l]}$ Output: 11$^{[LR, l]}$ 11$^{[LR, l]}$ = (17a,b)

	*LR$_{NEW}$-LR$^{[LOC1]}$	*31	ID-T	ID-LR	*11	*LR$_{OLD}$-LR$^{[LOC1]}$
☞a.11$^{[LR, l]}$ 11$^{[LR, l]}$					**	*
b.31$^{[HR, l]}$ 11$^{[LR, l]}$	*!	*	*	*	*	

Irrelevant constraints and candidates are omitted in (44) to avoid digression. In candidate (b), the violation locus bears a new LR, which violates the top-ranked *LR$_{NEW}$-LR$^{[LOC1]}$. In candidate (a), the violation locus bears an old LR, which violates only the bottom-ranked *LR$_{OLD}$-LR$^{[LOC1]}$. Candidate (a) thus emerges.

As discussed so far, both *LR-LR and *HR-HR are active in Zhuolan Raoping. *LR-LR disallows 31 to map to 11 before 31, 11, and 2, while *HR-HR prevents 31 from mapping to 33 before 53, 44 or 5. A question then arises from tri-tonal strings like [HR 31 LR], where *LR-LR and *HR-HR are bound to be in conflict. If the medial 31 maps to 11, the surface […LR LR] pair would violate *LR-LR; conversely, if the medial 31 maps to 33, the surface [HR HR…] pair would violate *HR-HR. In the corpus, there are 113 trisyllabic and tetrasyllabic tokens consisting of [HR 31 LR] in the input; all the medial 31's map to 33's, but not 11's. Some examples are given in (45).

(45) HR-31-LR → HR-33-LR

a. *hang* *zo* *ha* 'to enter the kitchen'
53 31 11
[HR, hl] [LR, h1] [LR, 1] input
53 33 11
[HR, l] [HR, 1] [LR, 1] output

b. *tai* *du* *shi* 'big stomach'
55 31 31
[HR, hl] [LR, h1] [LR, 1] input
33 33 31
[HR, hl] [HR, 1] [LR, h1] output

c. *shid* *lo* *gug* 'to eat old rice'
$\underline{5}$ 31 $\underline{2}$
[HR, h] [LR, h1] [LR, 1] input
$\underline{2}$ 33 $\underline{2}$
[HR, h] [HR, 1] [LR, 1] output

The "31 → 33" mappings in (45) yield the surface [HR-33-LR] sequences, which sacrifice *HR-HR to satisfy *LR-LR. One may suggest that *HR-HR is ranked below *LR-LR. This is true in the corpus, where 53 and 44 map to 33 before a high register tone. However, the "*LR-LR >> *HR-HR" ranking does not account for the mappings in (46).

(46) LR-31-HR → LR-11-HR

a. *chi* *fi* *ngiu* 'to kill a buffalo'
11 31 53
[LR, l] [LR, h1] [HR, h1] input
11 11 53
[LR, l] [LR, 1] [HR, h1] output

b. *ten* *gieu* *pui* 'to hear dog's barking'
11 31 55
[LR, l] [LR, h1] [HR, h] input
11 11 55
[LR, l] [LR, 1] [HR, h] output

c. *dong* *pui* *hab* ‘very cooperative’

11 31 <u>5</u>

[LR, l] [LR, h1] [HR, h] input

11 11 <u>5</u>

[LR, l] [LR, 1] [HR, h] output

In the corpus, there are 249 trisyllabic and tetrasyllabic tokens that contain that the [11 31 HR] input, and 239 of the medial 31’s map to 11’s, but only 10 of them map to 33’s. The derived [11 11 HR] pattern ignores the violation of *LR-LR.

(47) Input: $55^{[HR,\ h]}$ $31^{[LR,\ hl]}$ $31^{[LR,\ hl]}$ Output: $55^{[HR,\ h]}$ $33^{[HR,\ l]}$ $31^{[LR,\ hl]}$ = (45b)

	$*LR_{NEW}$-$LR^{[LOC1]}$	*55 *31	ID-HR	*HR-HR	ID-LR	$*LR_{OLD}$-$LR^{[LOC1]}$
a. $55^{[HR,\ h]}$ $31^{[LR,\ hl]}$ $31^{[LR,\ hl]}$		***!				*
☞b. $55^{[HR,\ h]}$ $33^{[LR,\ hl]}$ $31^{[LR,\ hl]}$		**		*	**	
c. $55^{[HR,\ h]}$ $11^{[LR,\ hl]}$ $31^{[LR,\ hl]}$	*!	**		*	**	*

(48) Input: $11^{[LR,\ l]}$ $31^{[LR,\ hl]}$ $53^{[HR,\ hl]}$ Output: $11^{[LR,\ l]}$ $11^{[LR,\ l]}$ $53^{[HR,\ hl]}$ = (46a)

	$*LR_{NEW}$-$LR^{[LOC1]}$	*31 *53	ID-HR	*HR-HR	ID-LR	$*LR_{OLD}$-$LR^{[LOC1]}$
a. $11^{[LR,\ l]}$ $31^{[LR,\ hl]}$ $53^{[HR,\ hl]}$		**!			**	*
b. $11^{[LR,\ l]}$ $33^{[LR,\ hl]}$ $53^{[HR,\ hl]}$		*			*!	
☞c. $11^{[LR,\ l]}$ $11^{[LR,\ hl]}$ $53^{[HR,\ hl]}$		*				*

4. Tone Sandhi at the Syntax-phonology Interface

The syntax-phonology interface has been a focus of attention for over three decades. The standard prosodic theory posits prosodic structure as an intermediate level between syntax and phonology. On the one hand, the prosodic hierarchy is layered downward from the utterance at one end and upward from the syllable at the other. On the other hand, the prosodic constituents are derived on the basis of syntactic information, relation-based or end-based. The relation-based parameters are developed in works by Nespor & Vogel (1986), Shih (1986), Hung (1987), Z. Zhang (1988), Bickmore (1989), Inkelas (1989), Kanerva (1989) and H. Zhang (1992), among others. These parameters require prosodic constituents to be built on syntactic relations, such as immediate constituency, c-command, m-command and the like. The end-based parameters are developed in works by Selkirk (1986), Chen (1987), L. Cheng (1987), Selkirk & Shen (1990), Hsiao (1991, 1995), and H. Hsu (1994), among others. These parameters locate designated edges of syntactic projections, XP or X^{head}, to mark boundaries of prosodic structures. In either case, a derived prosodic constituent serves as a rule domain to confine phonological rules, tonal, accentual or segmental.

4.1 Tonal Phrasing

There are 3212 trisyllabic tokens in the corpus, which consists of 14 types of syntactic structures, including [V [NN]^{NP}]^{VP}, [Av V N]^{VP}, [VV V]^{VP}, [Av AN]^{AP}, [[Av Av]^{AdvP} V]^{VP}, [V [N]^{NP} V]^{VP}, [A NN]^{NP}, [N NN]^{NP}, [NN N]^{NP}, [NN CL]^{NP}, [[NN]^{NP} V]^{S}, [[N]^{NP} [VR] ^{VP}]^{S}, [[N]^{NP} [Av A]^{AP}]^{S}, and [[N]^{NP} [V [N]^{NP}]^{VP}]^{S}. The tones in the corpus are coded with several notations: B represents base tone and S represents sandhi tone;]# indicates a line-final position;]^{NP}-,]^{VP}-,]^{AvP}- and the like indicates a phrase-final but not line-final position; σ indicates syllable; B- and S- indicate projection-nonfinal tones. Due to limited space, several abbreviations are used in the following discussions: Av/AvP stands for an adverb/adverbial phrase, CL for a classifier, LOC for a localizer, SUF for a suffix, and PA for a particle. The occurrences of base tones and sandhi in the trisyllabic structures are summarized as follows.

(49) Tone Distribution in Trisyllabic Structures

	B]#	S]#	B]^{NP}-	S]^{NP}-	B]^{AvP}-	S]^{AvP}-	B-	S-	σ	Lines
[V [NN]^{NP}]^{VP}	167	0					10 (52)	272	501	167
[Av V N]^{VP}	279	0					12 (90)	456	837	279
[VV V]^{VP}	142	0					7 (30)	247	426	142
[Av AN]^{AP}	280	0					14 (34)	512	840	280
[[Av Av]^{AvP} V]^{VP}	192	0			2 (23)	167	0 (14)	178	572	192
[V [N]^{NP} V]^{VP}	350	0	339	11			0 (22)	328	1050	350
[A NN]^{NP}	312	0					5 (54)	565	936	312
,[N NN]^{NP}	230	0					0 (19)	441	690	230
[NN N]^{NP}	340	0					3 (60)	617	1020	340
[NN CL]^{NP}	160	0					2 (18)	300	580	160

[[NN]^{NP} V]^{S}	125	0	121	4			10 (11)	104	375	125
[[N]^{NP} [VR]^{VP}]^{S}	189	0	165	24			0 (36)	153	567	189
[[N]^{NP} [Av A]^{AP}]^{S}	148	0	112	36			0 (44)	104	444	148
[[N]^{NP} [V [N]^{NP}]^{VP}]^{S}	298	0	253	45			3 (68)	227	894	298
Total	3212	0	990	120	2 (23)	167	66 (552)	4504	9636	3212

The parentheses in the "B-" column means that among the projection-nonfinal base tones the parenthesized number of tones are Yin Ping tones, 11's, which are not subject to tone sandhi in this Hakka dialect.

There are 6122 tetrasyllabic tokens in the corpus. The tetrasyllabic tokens display 26 types of syntactic structures, including [V [V [NN]^{NP}]^{VP}]^{VP}, [Av [V [NN]^{NP}]^{VP}]^{VP}, [[NN N]^{NP} [V]^{VP}]^{S}, [V [NN N]^{NP}]^{VP}, [[V [NN]^{NP}]^{VP} V]^{VP}, [[NN]^{NP} [Av V]^{VP}]^{S}, [[NN]^{NP} [Av A]^{AP}]^{S}, [[VN]^{N} [NN]^{N}]^{NP}, [[NN N]^{NP} LOC]^{NP}, [VV [NN]^{NP}]^{VP}, [V [AN N]^{NP}]^{VP}, [[NN N]^{N} N]^{NP}, [[[N]^{NP} [V [NN]^{NP}]^{VP}]^{S}, [[Av [NN]^{NP}]^{AvP} V]^{VP}, [[V [NN]^{NP}]^{AvP} V]^{VP}, [[Av [VN V]^{AP}]^{AvP}, [V [[Av Av]^{AvP} V]^{VP}]^{VP}, [[VV [Av V]^{VP}]^{VP}, [[Av V [VV]^{NP}]^{VP}, [[AA N]^{NP} V]^{VP}, [[V [NN SUF]^{NP}]^{VP}, [[V [NN]^{NP}]^{VP} PA]^{s}, [[N]^{NP} [V [NN]^{NP}]^{VP}]^{S}, and [[N]^{NP}[V[Av Av]^{AvP}]^{VP}]^{S}. The occurrences of base tones and sandhi in the tetrasyllabic structures are summarized as follows.

(52) Tone Distribution in Tetrasyllabic Structures

	B]#	S]#	B]^{NP}-	S]^{NP}-	B]^{AvP}-	S]^{AvP}-	B-	S-	σ	Lines
[V [V [NN]^{NP}]^{VP}]^{VP}	200	0					0 (54)	546	800	200
[Av [V [NN]^{NP}]^{VP}]^{VP}	335	0					0 (11)	994	1340	335
[[NN N]^{NP} [V]^{VP}]^{S}	320	0	298	22			0 (9)	631	1280	320
[V [NN N]^{NP}]^{VP}	402	0					7 (67)	1132	1608	402
[[V [NN]^{NP}]^{VP} V]^{VP}	133	0	130	3			1 (16)	249	532	133
[[NN]^{NP} [Av V]^{VP}]^{S}	105	0	104	1			0 (2)	208	420	105

[[NN]^{NP} [Av A]^{AP}]^{S}	99	0	90	9			3 (8)	187	396	99
[[VN]^{N} [NN]^{N}]^{NP}	235	0					21 (92)	592	940	235
[[NN N]^{NP} LOC]^{NP}	220	0	30	190			0 (23)	417	880	220
[VV [NN]^{NP}]^{VP}	256	0					15 (61)	692	1024	256
[V [AN N]^{NP}]^{VP}	135	0					0 (23)	382	540	135
[V [A NN]^{NP}]^{VP}	201	0					0 (31)	572	804	201
[[[N]^{NP} [V [NN]^{NP}]^{VP}]^{S}	435	0	434	1			0 (13)	857	1740	435
[[NN N]^{N} N]^{NP}	467	0					5 (46)	1350	1868	467
[[Av [NN]^{NP}]^{AvP} V]^{VP}	120	0	100	20			10 (33)	197	480	120
[[V [NN]^{NP}]^{AvP} V]^{VP}	269	0	259	10			2 (98)	438	1076	269
[V [[Av Av]^{AvP} V]^{VP}]^{VP}	215	0			0	215	0 (19)	411	860	215
[[Av [VN V]^{AP}]^{AvP}	112	0					0 (9)	327	448	112
[[V [Av Av]^{AvP} V]^{VP}	190	0			0	190	0 (11)	369	760	190
[[VV [Av V]^{VP}]^{VP}	343	0					3 (39)	987	1372	343
[[Av V [VV]^{NP}]^{VP}	233	0					0 (55)	644	932	233
[[AA N]^{NP} V]^{VP}	330	0	299	31			19 (73)	568	1320	330
[[V [NN SUF]^{NP}]^{VP}	89	0					0 (12)	255	356	89
[[V [NN]^{NP}]^{VP} PA]^{s}	0	98	98	0			0 (8)	188	392	98
[[N]^{NP} [V [NN]^{NP}]^{VP}]^{S}	245	0	455	35			0 (5)	240	980	245

[[N]^{NP}[V[Av Av]^{AvP}]^{VP}]^{S}	335	0	312	23			0 (29)	641	1340	335
Total	6024	98	2609	345	0	405	86 (847)	14074	24488	6122

The statistic numbers thus suggest that the trisyllabic and tetrasyllabic tone sandhi tends to be blocked at phrase-final and line-final positions. Like Taiwanese (Chen 1987, 2000, Hsiao 2000), the tone sandhi domain in Zhuolan Raoping is marked at the right edge of S, NP, VP and nonadjunct AP, but not AvP or adjunct AP. Cross-categorically, the tone sandhi domain is marked at the right edge of XP.

4.2 An OT Analysis of Tonal Phrasing

In terms of classic OT (Prince & Smolensky 1993/2004), the end-based approach (Selkirk 1986) is reframed under the notion of generalized alignment (McCarthy & Prince 1993). A prosodic domain often results from a nonsyntactic span due to the avoidance of recursion, and this fact constitutes the core argument for an independent prosodic constituent. Based on the corpus, the tone sandhi domain in Zhuolan Raoping is built by aligning the right edge of XP with the right edge of a prosodic constituent, in particular, the phonological phrase. Three constraints are proposed below.

(53) ALIGN-R(XP, φ): assign one violation mark for every XP whose right edge is not aligned with the right edge of a phonological phrase, φ. (XP ≠ adjunct)

(54) ALIGN-R(φ, XP): assign one violation mark for every phonological phrase, φ, whose right edge is not aligned with the right edge of an XP. (XP ≠ adjunct)

(55) Parse(ω): assign one violation mark for every phonological word, ω, that is not parsed by a phonological phrase, φ.

The alignment constraints are dominated by Parse(ω), which bans any extraprosodic element at the level of phonological phrase, φ.

(56) [[V [NN]^{NP}]^{VP} V]^{VP}

	Parse(ω)	ALIGN-R(φ,XP,)	ALIGN-R(XP, φ)
a. φ(V NN V)φ			*!
b. φ(V NN)φ V	*!		
☞c. φ(V NN)φ φ(V)φ			
d. φ(V)φ φ(NN V)φ		*!	*
e. φ(V)φ φ(NN)φ φ(V)φ		*!	

In (56), candidate (a), where the right edge of NP is not aligned with that of a φ, is ruled out by ALIGN-R(XP, φ), while candidate (b) by Parse(ω), since the final V is not parsed. The parenthesized “!” indicates that candidate (d) fatally violates either ALIGN-R(XP, φ), as NP is not right-aligned with a φ, or ALIGN-R(φ, XP), as the first φ is right-aligned with V, and the latter rules out candidate (e) for the same reason. Eventually, candidate (c) emerges, and the tetrasyllabic line is parsed into two phonological phrases.

(57) $[\mathrm{V}\ [[\mathrm{Av\ Av}]^{\mathrm{AvP}}\ \mathrm{V}]^{\mathrm{VP}}]^{\mathrm{VP}}$

	Parse(ω)	ALIGN-R(φ,XP,)	ALIGN-R(XP, φ)
☞a. φ(V Av Av V)φ			
b. φ(V Av Av)φ V	*!	*	
c. φ(V Av Av)φ φ(V)φ		*!	
d. φ(V)φ φ(Av Av V)φ		*!	
e. φ(V)φ φ(Av Av)φ φ(V)φ		*!*	

In (57), the AvP is an adjunct, which cannot be right-aligned with a φ. For this reason, ALIGN-R(φ, XP) rules out candidates (c-d). The same constraint rules out candidate (e), where the first V forms an φ. Candidate (b) is ruled out by Parse(ω), and thus candidate (a) is selected as the optimal output.

(58) $[[\mathrm{NN\ N}]^{\mathrm{NP}}\ [\mathrm{V}]^{\mathrm{VP}}]^{\mathrm{S}}$

	Parse(ω)	ALIGN-R(φ,XP,)	ALIGN-R(XP, φ)
a. φ(NN N V)φ			*!
b. φ(NN N)φ V	*!		
☞c. φ(NN N)φ φ(V)φ			
d. φ(NN)φ φ(N)φ φ(V)φ		*!	

(59) $[[\mathrm{NN\ N}]^{\mathrm{NP}}\ \mathrm{LOC}]^{\mathrm{NP}}$

	Parse(ω)	ALIGN-R(φ,XP,)	ALIGN-R(XP, φ)
*☞a. φ(NN N LOC)φ			*!
b. φ(NN N)φ LOC	*!		
👎c. φ(NN N)φ φ(LOC)φ			
d. φ(NN)φ φ(N)φ φ(LOC)φ		*!	

The three constraints in (53-55) correctly select candidate (c) in (58) as the optimal output, but make wrong predictions in (59), where candidate (c) is wrongly selected, as indicated by the 👎 symbol, but the real optimal output is candidate (a), which is wrongly ruled out by ALIGN-R(XP, φ), as indicated by the *☞ symbol. The contrast between (58) and (59) lies in the monosyllabic localizer, which is a prosodic clitic. In other words, a clitic does not form a phonological phrase alone. A constraint like (60) is needed then.

(60) *φ(*c*): assign one violation mark for every phonological phrase, φ, that is equal to a single monosyllabic clitic, *c*.

The constraint *φ(*c*) disallows any monosyllabic clitic to form a φ, and it dominates the alignment constraints, as in (61-62).

(61) [[NN N]^{NP} [V]^{VP}]^{S}

	Parse(ω)	*φ(*c*)	ALIGN-R(φ,XP)	ALIGN-R(XP, φ)
a. φ(NN N V)φ				*!
b. φ(NN N)φ V	*!			
☞c. φ(NN N)φ φ(V)φ				
d. φ(NN)φ φ(N)φ φ(V)φ			*!	

(62) [[NN N]^{NP} LOC]^{NP}

	Parse(ω)	*φ(*c*)	ALIGN-R(φ,XP)	ALIGN-R(XP, φ)
☞a. φ(NN N LOC)φ				
b. φ(NN N)φ LOC	*!			
c. φ(NN N)φ φ(LOC)φ		*!		
d. φ(NN)φ φ(N)φ φ(LOC)φ			*!	

In (62), the monosyllabic locative forms a phonological phrase, and thus is expectedly ruled out by *φ(*c*).

Selkirk (2011) rethinks her (1986) end-parameters and raises a doubt on whether a nonsyntactic span really forms a prosodic constituent. She suggests that further empirical evidence is required, and she proposes a theory of "match" that indicates a tendency for prosodic constituents to mirror syntactic constituents. In this sense, the syntax-prosody match is no longer a matter of alignment (or markedness) but operates on correspondence (or faithfulness). In this sense, the alignment constraints can be replaced by the match constraint in (63).

(63) MATCH(XP, φ): assign one violation mark for every XP that is not matched by a corresponding phonological phrase, φ. (XP ≠ adjunct)

However, the match constraint renders incorrect predictions in the following tableaux.

(64) [[V [NN]^NP]^VP V]^VP

		Parse(ω)	*φ(*c*)	MATCH (φ,XP)	MATCH (XP, φ)
☜a.	φ(φ(V φ(NN)φ)φ V)φ				
b.	φ(V φ(NN)φ)φ V	*!			*
*☞c.	φ(V φ(NN)φ)φ φ(V)φ			*!	
d.	φ(V)φ φ(φ(NN)φ V)φ			*!*	*
e.	φ(V)φ φ(NN)φ φ(V)φ			*!*	**

(65) [[NN N]^NP LOC]^NP

		Parse(ω)	*φ(*c*)	MATCH (φ,XP)	MATCH (XP, φ)
☜a.	φ(φ(NN N)φ LOC)φ				
b.	φ(NN N)φ LOC	*!			*
c.	φ(NN N)φ φ(LOC)φ		*!	*	*
d.	φ(NN)φ φ(N)φ φ(LOC)φ			*!*	***
*☞e.	φ(NN N LOC)φ				*!

In (64), candidate (a) is incorrectly selected, as it violates no constraint. Candidate (b) is ruled out by Parse(ω). Candidate (c) is the actual optimal output but is ruled out by MATCH (φ,XP), since the rightmost φ matches V but not VP, or any XP. Both candidate (d) and candidate (e) violate MATCH (φ,XP) twice, and thus both are ruled out. In (65), the actual optimal output is candidate (e), but it is ruled out by MATCH (XP, φ), since the inner NP does not match a phonological phrase. It should be noted that *φ(*c*) is unable to rule out candidate (c) here, since the outer VP matches the outer φ. In addition to the incorrect predictions, two theoretical consequences of the match theory are in order. First, the strict layer hypothesis, which prohibits recursion of prosodic category, is abandoned. Second, prosodic constituents are totally identical to (or mirror) syntactic constituents, in which case, the motivation for positing an independent prosodic level is unclear.

5. Summary

This research project has achieved several tasks. First, I establish a large corpus of Zhuolan Raoping, which contains 10662 tokens, monosyllabic to tetrasyllabic. The corpus is constructed with phonological coding, such as syllables, tones, syntactic and prosodic structures. Second, based on the corpus, I have some fresh observations on the tone sandhi of this Hakka dialect. i) Three smooth tones, 31, 44 and 53, are neutralized as either 33 or 11 in nonfinal position, and the tone neutralization is triggered by a set of tone markedness constraints. ii) The emergence of 33 or 11 is determined by the OCP effects on the registers.

iii) Tonal variation is keyed to the interaction between register OCP and register identity. At this point, stochastic evaluation is able to render predictions of particular frequency and language change. Finally, I posit a set of constraints to govern the formation of tone sandhi domain, and compare the standard prosodic theory and match theory. I have shown that the match constraints are not well-motivated, and may make incorrect predictions. To further investigate the relevant phonological phrasing and intonational phrasing, we will need to look at longer phrases and sentences. In the following project (MOST# 105-2410-H-004-181 -MY2), I am engaged in the construction of a corpus containing pentasyllabic and polysyllabic structures.

References

Anttila, Arto. 1997. Deriving variation from grammar. In Hinskens, Frans, Roeland van Hout, and Leo Wetzels Variation (Eds.), *Change and Phonological Theory* (pp. 35-68). Amsterdam: John Benjamins.

Bao, Zhi-ming. 1990. On the nature of tone. Ph. D. dissertation. Massachusetts Institute of Technology.

Bao, Zhi-ming. 1999. *The structure of tone*. Oxford, UK: Oxford University Press.

Bickmore, Lee. 1989. Kinyambo Prosody. Doctoral dissertation. University of California, Los Angeles.

Boersma, Paul. 1997. How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 21:43-58.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interactions Between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics. [University of Amsterdam dissertation.]

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32:45-86.

Coetzee, Andries W. 2006. Variation as accessing 'non-optimal' candidates. *Phonology* 23:337-385.

Chao, Yuen-ren. 1930. A system of tone letters. *Le Maˆitre Phone´tique* 45:24–47.

Chen, Matthew Y. 1987. The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4:109-150.

Chen, Matthew Y. 2000. *Tone Sandhi: Patterns across Chinese Dialects*. Cambridge University Press, Cambridge.

Cheng, Li-shan. 1987. Derived domains and Mandarin third tone sandhi. *Chicago Linguistic Society* 23: 16-29.

Chung, Raung-fu, 1996. The Segmental Phonology of Southern Min in Taiwan, Crane Publishing Co., Ltd., Taipei.

Duanmu, San. 1990. A Formal Study of Syllable, Tone, Stress and Domain in Chinese

Languages. Ph. D. Dissertation. MIT.
Hayes, Bruce. 1989. The prosodic hierarchy in meter. *Phonetics and Phonology* 1: 201-260.
Hsiao, Yuchau E. 1991. *Syntax, Rhythm and Tone: A Triangular Relationship.* Doctoral Dissertation. University of California, San Diego. Taipei: Crane Publishing Co., Ltd..
Hsiao, Yuchau E. 1995. Southern Min Tone Sandhi and Theories of Prosodic Phonology. Student Book Co., Ltd., Taipei.
Huang, Tzu-chuan. 2011. An OT Approach to the Tone Sandhi of Checked Syllables in Shanghai. M.A. Thesis. National Chengchi University.
Huang, Tzu-chuan, and Yuchau E. Hsiao. 2016. Analogical leveling of tonal paradigms in two Hakka dialects. The 11th International Symposium on Taiwanese Languages and Teaching. Taipei: Academia Sinica.
Hyman, Larry. 1993. Register tones and tonal geometry. In *The Phonology of Tone: The Representation of Tonal Register*, eds. K. Snider and H. van der Hulst, 75-108. Berlin: Mouton de Gruyter.
Hsu, Hui-chuan. 1994. Constraint-Based Phonology and Morphology: A Survey of Languages in China. Doctoral Dissertation. University of California, San Diego.
Hung, Tony. 1987. Syntactic and Semantic Aspects of Chinese Tone Sandhi. Doctoral dissertation, University of California, San Diego.
Inkelas, Sharon. 1987. Tone feature geometry,. *Northeast Linguistic Society* 18: 222-237.
Inkelas, Sharon. 1989. Prosodic Constituency in the Lexicon. Doctoral dissertation. Stanford University.
Inkelas, Sharon, and Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45.1:133-171.
Kanerva, Jonni. 1989. Focus and Phrasing in Chichewa Phonology. Doctoral dissertation. Stanford University.
Li, Zhiqiang. 2003. The Phonetics and Phonology of Tone Mapping in a Constrain-based Approach. Ph. D. Dissertation. MIT.
Lin, Yen-Hwei. 2007. *The sounds of Chinese*. Cambridge University Press, Cambridge, U.K., New York, U.S.A.
Lu, Mingchang. 2013. Modeling salience and prosody in loanword adaptation: cases of English [r] in Mandarin. *Concentric: Studies in Linguistics* 32 (2): 1-32.
Lu, Mingchang. 2016. Where Variation Vanishes: Constructing the Stochastic Perception Grammar for English Loanwords in Taiwan Mandarin. Ph. D. Dissertation. National Chengchi University.
McCarthy, John. 2003. Comparative markedness. Theoretical Linguistics 29: 1-51.
Nespor, Marina and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris Publications.
Ping, Yueling. 2001. Shanghai shiqu fangyan shengdiao shiyan yanjiu [Experimental studies

of Shanghai Chinese tonal system]. In *Wuyu Shengdiao de Shiyan Yanjiu* [*Experimental studies of Wu Chinese Tonal systems*], eds. by Yang Jianqiao, and You Rujie 17-38. Shanghai: Shanghai Education Publishing House.

Prince, Alan and Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar.* Blackwell, Malden, Massachusetts, and Oxford, UK. Revision of 1993 technical report. ROA-537.

Qian, Nairung. 1992. *Dangdai Wuyu Yanjiu* [*Studies in Contemporary Wu Dialects*]. Shanghai: Shanghai Education Press.

Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3:371-405.

Selkirk, Elizabeth, 2011. The Syntax-phonology interface. Goldsmith. J., Riggle, J., Yu, A. eds., *The Handbook of Phonological Theory*, 2nd edition. Oxford: Blackwell Publishing, pp. 435-484.

Selkirk, Elizabeth and Tong Shen. 1990. Prosodic Domains in Shanghai Chinese. S. Inkelas, and D. Zec, eds., *The Phonology-Syntax Connection*. Chicago: University of Chicago Press, pp. 313-338.

Shih, Chilin. 1986. The Prosodic Domain of Tone Sandhi in Chinese. Doctoral dissertation. University of California, San Diego.

Tesar, Bruce, and Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29:229–268.

Xu, Baohua, Tang Zhenzhu, You Rujie, Qian Nairong, Shi Rujie, and Shen Yaming. 1988. *Shanghai Shiqu Fangyanzhi [Urban Shanghai Dialects]*. Shanghai: Shanghai Jiaoyu Chupanshe [Shanghai Education Press].

Yip, Moria. 1980. The Tonal Phonology of Chinese. Ph. D. Dissertation. MIT.

Yip, Moria. 1989. Feature Geometry and Co-occurrence Restrictions. *Phonology* 6.2: 349-374.

Yip, Moria. 1993. Tonal Register in East Asian Languages. In *The Phonology of Tone: The Representation of Tonal Register*, eds. by H. van der Hulst and K. Snider, 245-268. Mouton de Gruyter, Berlin, NewYork.

Yip, Moira. 2001. Tonal features, tonal inventories and phonetic targets. In *UCL Working Papers in Linguistics,* 161-188.

Yip, Moria. 2002. *Tone*. Cambridge: Cambridge University Press.

Zee, Eric, and Ian Maddieson. 1979. Tones and Tone Sandhi in Shanghai: Phonetic Evidence and Phonological Analysis. *UCLA Working Papers in Phonetics* 45: 93-129.

Zhang, Hongming. 1992. Topics in Chinese Phrasal Phonology. Doctoral dissertation. University of California, San Diego.

Zhang, Zhengsheng. 1988. Tone and Tone Sandhi in Chinese. Ph.D. Dissertation. The Ohio State University.

呂嵩雁(Lu)。1993。《臺灣饒平方言》。臺北市：東吳大學中國文學研究所碩士論文。
徐貴榮(K. Hsu)。2008。《臺灣饒平客話的源與變》。新竹：國立新竹教育大學臺灣語言與語文教育研究所博士論文。
徐瑞珠(R. Hsu)。2005。《苗栗卓蘭客家話研究》。高雄：國立高雄師範大學臺灣語言及教學研究所碩士論文。
涂春景(Tu)。1998。《苗栗卓蘭客家方言詞彙對照》。臺北：撰者自印本。
彭美慈(Peng)。2007 。《臺灣卓蘭饒平客語音韻研究》。臺灣大學碩士論文。
鄭明中、徐嘉駿(M. Cheng & J. Hsu)。2013。〈卓蘭鎮饒平客家話聲調之聲學分析〉。第一屆台灣客家語言及其教學研討會論文。新竹：國立新竹教育大學臺灣語言與語文教育研究所。
羅肇錦(Luo)。2007。《重修苗栗縣志－語言志》。苗栗：苗栗縣政府。

附件五

科技部補助專題研究計畫出席國際學術會議心得報告

日期：2017 年 2 月 28 日

計畫編號	#MOST-103-2410-H-004-086-MY2		
計畫名稱	從句法音韻介面觀點研究卓蘭饒平三字組及四字組之連讀變調 2/2		
出國人員姓名	蕭宇超	服務機構及職稱	國立政治大學語言學研究所教授
會議時間	2016 年 8 月 10 日至 2016 年 8 月 13 日	會議地點	Rice University, Houston, USA
會議名稱	(英文) The 2nd International Conference on Spoken Chinese Corpora From Theory to Pedagogical Applications		
發表題目	(英文) A Corpus Analysis of Marking Influences on Accented L2		

2016 年 8 月 10~13 日赴美國休斯頓參加「The 2nd International Conference on Spoken Chinese Corpora From Theory to Pedagogical Applications」，發表論文題目為：A Corpus Analysis of Marking Influences on Accented L2。承蒙科技部研究計畫補助旅費，特此感謝。此外，也要感謝 Rice Universit 葉萌教授的接待。此次會議出席者踴躍，約有百餘人，其中不乏知名學者，也有不少年輕學者及研究生與會，值得高興的是，我看到台灣政治大學的教授及研究生出席會議，並宣讀論文。此次會議安排發表的論文有數十篇，在語料庫方面的論文則有 10 篇，會議語料庫論文詳列如下：

Hongyin Tao (陶红印), UCLAExploiting Spoken Corpora for Chinese Language Research and Teaching Elastic word length in Old Chinese and Middle Chinese.

Baolin Zhang (張寶林), Beijing Language and Culture University. 為漢語課堂教學服務—漢語中介語語料庫建設與發展的新思路

Liang Fu (傅亮), Rice University. Using Chinese Spoken Corpus to Teach Chinese for Medical Purposes.

Yuchau Hsiao, National Chengchi University. A Corpus Analysis of Marking Influences on Accented L2.

Yuming Chen (陳玉明), Education University. 基於語料庫探討量詞「支」的泛化使用現象.

Shan He (何姗), Beijing Norman University, Iowa University. Chen Mengtian (陳夢恬), Iowa University. 淺談漢語口語語料库的語料選取問題.

Delfina Chuang (莊錦惠), El Salvador University. 從放大鏡到顯微鏡窺探語料庫的設計脈絡與使用心得.

Xianwen Cao (曹賢文), Nanjing University. 漢語二語習得研究與學習者口語語料庫建設.

Chin-Chin Tseng (曾金金), Taiwan Normal University. Building spoken interlanguage corpora from online interaction for L2 Chinese teaching and learning.

Meng Yeh (葉萌), Rice University. Teaching Interactional Practices through Building Corpora of Spoken Languages.

我的論文口頭報告安排於 8 月 10 日(週三)上午場，13:30~14:15 之間。此次科技部計畫主要是建立大型語料庫，從句法音韻介面觀點討論卓蘭饒平三字組及四字組之連讀變調。此番會議，會中有不少語料庫專家參與討論，著實獲益良多。

科技部補助計畫衍生研發成果推廣資料表

日期:2017/03/29

科技部補助計畫	計畫名稱: 從句法音韻介面觀點研究卓蘭饒平三字組及四字組之連讀變調
	計畫主持人: 蕭宇超
	計畫編號: 103-2410-H-004-086-MY2　　學門領域: 音韻學

無研發成果推廣資料

103年度專題研究計畫成果彙整表

計畫主持人：蕭宇超	計畫編號：103-2410-H-004-086-MY2
計畫名稱：從句法音韻介面觀點研究卓蘭饒平三字組及四字組之連讀變調	

成果項目				量化	單位	質化（說明：各成果項目請附佐證資料或細項說明，如期刊名稱、年份、卷期、起訖頁數、證號...等）
國內	學術性論文	期刊論文		1	篇	2015.10. Rethinking OCP effects on tone sandhi. Language and Linguistics. 16.6: 927－945. [SSCI, A&HCI].
國內	學術性論文	研討會論文		7	篇	2016/7. Disyllabic and trisyllabic tone sandhi in Zhuolan Raoping: a corpus analysis. IACL-24. 2016/5. Tonal variation in Zhuolan Raoping. ISCLL-15. 2016/5. Sandhi variation of the high-register tones in Raoping Hakka. NACCL-28. 2015/8. Locus of Violation in Zhuolan Raoping tone sandhi. ICSTLL-48. 2015/8. OCP effects on Zhuolan Raoping tone neutralization. Paper to be presented at IACL-23 & ISSKL-1. 2015/4. Tone sandhi in Zhuolan Raoping: a perspective from the syntax-phonology interface. NACCL-27.
國內	學術性論文	專書		0	本	
國內	學術性論文	專書論文		2	章	2015.11. Tone sandhi in Zhuolan Raoping: a perspective from the syntax-phonology interface. Proceedings of NACCL-27. 359-367. 2017, to appear. Sandhi variation of the high-register tones in Raoping Hakka. Proceedings of NACCL-28.
國內	學術性論文	技術報告		1	篇	Trisyllabic and tetrasyllabic tone sandhi in Zhuolan Raoping: a perspective from the syntax-phonology interface (MOST# 103-2410-H-004-086-MY2)
國內	學術性論文	其他		0	篇	
國內	智慧財產權及成果	專利權	發明專利 申請中	0	件	
國內	智慧財產權及成果	專利權	發明專利 已獲得	0	件	
國內	智慧財產權及成果	專利權	新型/設計專利	0	件	

		商標權		0		
		營業秘密		0		
		積體電路電路布局權		0		
		著作權		0		
		品種權		0		
		其他		0		
	技術移轉	件數		0	件	
		收入		0	千元	
國外	學術性論文	期刊論文		0	篇	
		研討會論文		0		
		專書		0	本	
		專書論文		0	章	
		技術報告		0	篇	
		其他		0	篇	
	智慧財產權及成果	專利權	發明專利 申請中	0	件	
			發明專利 已獲得	0		
			新型/設計專利	0		
		商標權		0		
		營業秘密		0		
		積體電路電路布局權		0		
		著作權		0		
		品種權		0		
		其他		0		
	技術移轉	件數		0	件	
		收入		0	千元	
參與計畫人力	本國籍	大專生		0	人次	
		碩士生		0		
		博士生		0		
		博士後研究員		0		
		專任助理		0		
	非本國籍	大專生		0		
		碩士生		0		
		博士生		0		
		博士後研究員		0		
		專任助理		0		
其他成果 （無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體						

效益事項等，請以文字敘述填列。）	

科技部補助專題研究計畫成果自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現（簡要敘述成果是否具有政策應用參考價值及具影響公共利益之重大發現）或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估 ■達成目標 □未達成目標（請說明，以100字為限） 　　□實驗失敗 　　□因故實驗中斷 　　□其他原因 說明：
2. 研究成果在學術期刊發表或申請專利等情形（請於其他欄註明專利及技轉之證號、合約、申請及洽談等詳細資訊） 論文：■已發表　□未發表之文稿　□撰寫中　□無 專利：□已獲得　□申請中　■無 技轉：□已技轉　□洽談中　■無 其他：（以200字為限） 本研究計畫建立了一個大型的卓蘭饒平語料庫，其中包含10662個單音節至四音節樣本，並建置各種音韻標注。此語料庫將可實質助益進一步研究。
3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性，以500字為限） 在理論上，本研究觀察到，卓蘭饒平的聲調中和由聲調標記、調域OCP及調域信實所啟動。此外，本研究比較「標準韻律理論」與「契合理論」，發現契合制約可能做出錯誤的預測。
4. 主要發現 本研究具有政策應用參考價值：■否　□是，建議提供機關 （勾選「是」者，請列舉建議可提供施政參考之業務主管機關） 本研究具影響公共利益之重大發現：■否　□是 說明：（以150字為限）