

MASTER'S THESIS

Expletive Epenthesis in Hong Kong Cantonese

TAM, Nga Wing Olivia

Date of Award:
2020

[Link to publication](#)

General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

**EXPLETIVE EPENTHESIS IN
HONG KONG CANTONESE**

TAM NGA WING, OLIVIA
STUDENT NO. 19442866

A Dissertation Submitted In Partial
Fulfilment Of The

MASTER OF ARTS IN LANGUAGE STUDIES

HONG KONG BAPTIST UNIVERSITY

1 DECEMBER 2020

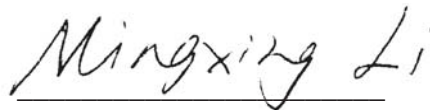
HONG KONG BAPTIST UNIVERSITY

1 December 2020

We hereby recommend that the Dissertation by Ms. Tam Nga Wing, Olivia entitled “Expletive Epenthesis in Hong Kong Cantonese” be accepted in partial fulfilment of the requirements for the degree of Master of Arts in Language Studies.



Prof. Lian-Hee Wee
Principal Supervisor



Dr. Mingxing Li
Second Reader

Acknowledgements

I wish to thank my supervisor, Prof. Lian-Hee Wee, for his suggestions and feedback throughout the course of work on this dissertation. Since we met the first time six years ago, Lian-Hee has always given me inspiration and support. His lessons on linguistics and creative ways of thinking are the fruitful time I spent in university that I will never forget. His teaching and advice are critical to the completion of this dissertation and success in my undergraduate and postgraduate years. He makes learning linguistics meaningful and enjoyable and being in the field has become my dream ever since.

I would like to express my gratitude to Dr. Mingxing Li for his trust. It has been one useful and wonderful experience working as a research assistant for Dr. Li since it is where I put my knowledge to practice and gain valuable learning opportunities and acquire practical skills in HKBU Phonology Laboratory.

Special thanks to everyone I have met in Hong Kong Baptist University and I am sure that I cannot make it this far without the help from the university and the English Department. I have met great teachers and friends in this journey that has allowed me to grow and spend quality time together. I would in particular hope to thank Vesper Liu, Kaixin Zhao, Dr. Suying Yang, Dr. John Wakefield, Dr. Janice Wong and Prof. Tony Hung for meaningful discussions and learning.

Finally, I wish to thank and express my love to my parents Liza Kwok and Philip Tam for their unlimited support and unconditioned love all the time.

I note that the entire dissertation is my own work and carried out by myself. It has been a hard year for all of us and I hope the pandemic will pass soon so we can get hold of our loved ones and continue to work harder towards our dreams.



Tam Nga Wing, Olivia

M.A. in Language Studies
Hong Kong Baptist University

Date: 1 December 2020

Total number of words: 10250

Abstract Monosyllabic expletives (curse words) like *diu2*, *lan2*, *kau1*, *ceon1* and *hai1* are often inserted within utterances in daily colloquial conversations of Hong Kong Cantonese and there seems to be some patterns to the way they are used. For instance, among the five available slots we can put to ‘Ma-lay-si-a’, prefixing or suffixing the monosyllabic expletive to the quadrisyllabic translated name is not possible and the position between -si- and -a is not acceptable. We would like to know why is there such a preference to the insertion places. This dissertation aims to investigate Hong Kong Cantonese expletive epenthesis within the framework of Optimality Theory. Through analysing monomorphemes and polymorphemes with two to five syllables, the author provides an Optimality Theory analysis to explain the puzzle is the result of complex interactions of prosody, foot and syntax and can predict where the expletive can appear with five constraints.

Keywords Hong Kong Cantonese · Optimality Theory · Prosody · Alignment

Table of Contents

Page of Acceptance	<i>i</i>
Acknowledgements	<i>ii</i>
Abstract	<i>iii</i>
Table of Contents	<i>iv</i>
Lists of items	<i>v</i>
1. Introduction	1
1.1 The puzzle	1
1.2 Research Questions	2
1.3 Structure of the dissertation	4
2. Background	5
3. Observations	7
3.1 Monomorphemes	7
3.2 Polymorphemes	11
4. Previous studies	15
5. Theoretical Framework	20
5.1 Optimality Theory	20
5.2 Relevant constraints	22
6. OT analysis and constraints ranking	26
6.1 Monomorphemes	26
6.2 Polymorphemes	30
7. Conclusion	39
References	40

List of items

- (1) Central Research Question
- (2) E in disyllabic monomorphemes
- (3) E in trisyllabic monomorphemes
- (4) E in quadrisyllabic monomorphemes
- (5) E in monomorphemes with five syllables
- (6) Patterns of expletive epenthesis in monomorphemes
- (7) E in disyllabic polymorphemes
- (8) E in left-branching (2+1) trisyllabic polymorphemes
- (9) E in right-branching (1+2) trisyllabic polymorphemes
- (10) E in quadrisyllabic polymorphemes
- (11) Overall observations of expletive epenthesis in Hong Kong Cantonese
- (12) Constraints (Fun, 2005)
- (13) Optimality Theory framework
- (14) Optimality Theory assessment tableau
- (15) Faithfulness constraint
- (16) Constraints on foot structure
- (17) FT BIN >> PARSE-SYLL, from /bicara/ ‘speak’
- (18) Alignment constraint
- (19) Generalised Alignment
- (20) Optimal candidate for E + *ping4 gwo2* ‘apple’
- (21) Optimal candidate for E + *si6 baa1 naa4* ‘spanner’
- (22) Optimal candidate for E + *maa5 loi6 sai1 aa3* ‘Malaysia’

- (23) Tree diagrams for *sik6 faan6* ‘eat rice’ and *nei5 sik6* ‘you eat’
- (24) Optimal candidate for E + *sik6 faan6* ‘eat rice’
- (25) Optimal candidate for E + *ci1 sin3* ‘crazy’
- (26) ALIGN EDGE >> FF >> FT BIN
- (27) Tree diagrams for *ping4 gwo2 syu6* and *hung4 ping4 gwo2*
- (28) Optimal candidate for E + *ping4 gwo2 syu6* ‘apple tree’
- (29) Optimal candidate for E + *hung4 ping4 gwo2* ‘red apple’
- (30) Optimal candidate for E + *caai4 ho2 fu1 si1 gei1* ‘Tchaikovsky’
- (31) Optimal candidate for E + *zing2 ping4 gwo2 pai1* ‘make apple pie’
- (32) Optimal candidate for E + *san1 sin1 ping4 gwo2* ‘fresh apple’
- (33) Optimal candidate for E + *fei1 leot6 ban1 jan4* ‘Filipino’
- (34) Hesse Diagram of the five constraints

Chapter One

Introduction

1.1 The puzzle

In colloquial Cantonese conversations, some Hong Kong people would insert curse words in their speech to intensify meanings or when they are emotional. *lan2* and *kau1* meaning ‘male genitalia’, *kwai2* ‘ghost’ and *hai1* ‘female genitalia’ are some common monosyllabic expletives used by many Hongkongers. These monosyllabic expletives can only be inserted within an utterance and they cannot come at the beginning or at the end of an utterance. Similar to the expletive infixation phenomenon in English, where native English speakers insert *bloody* into *California*, to become *Cali-bloody-fornia* (Hammond, 1999), expletive epenthesis in Hong Kong Cantonese (HKC) appears also to be infixal.

There are different patterns in terms of insertion in different kinds of strings in Hong Kong Cantonese, and the acceptable places for insertion is not solely determined by the syntactic boundary. For instance, when a monosyllabic expletive is inserted in 2+1 constructions like *ping4 gwo2 syu6* ‘apple tree’, it has to be *ping4 E gwo2 syu6* (in which E represents the monosyllabic expletive), but not **ping4 gwo2 E syu6* (in which * indicates unacceptability), suggesting that the insertion is not at where the syntactic boundary is. For 1+2 constructions like *daai6 ping4 gwo2* ‘big apple’, it is acceptable to say both *daai6 E ping4 gwo2* and *daai6 ping4 E gwo2*, where it becomes acceptable to fill the two available insertion slots with the expletive. In other constructions which has the same number of syllables and a similar internal syntactic structure, the insertion

place for the expletive, however, is not essentially the same. Taking a look at disyllabic strings like *chi1 sin3* ‘crazy’ (lit. cross wires), E can be inserted in the middle to become *chi1 E sin3*. For *jat1 ji6* ‘one two’, *jat1 E ji6* is not possible. The nature of the string and the relationship among the syllables may be factors affecting the possible insertion. The above examples show that E does not necessarily appear at syntactic boundaries and different internal structures of strings and relationships among syllables will affect the license of expletive epenthesis.

This dissertation aims to find out the possible insertion places of E in different kinds of strings and its pattern in Hong Kong Cantonese. The pattern will help to give a sense of what the prosodic system of native Hong Kong Cantonese speakers is like and leads us to finding principles to predict the acceptable insertion positions of E in any kinds of strings in the language.

1.2 Research Questions

Regarding patterns found in expletive epenthesis in Hong Kong Cantonese, there are three types of puzzles awaiting answers. Firstly, E cannot be put before or after an utterance or a string. This pattern may seem odd to some of us as Hong Kong Cantonese, like other Chinese languages, are syllable-timed languages in which insertion of some syllables would be expected to be rather free at the boundary of utterances. In other words, we would want to know why expletive epenthesis is infixal in Hong Kong Cantonese, which is similar to expletive infixation in American English (Hammond, 1999) but different from that phenomenon in Singaporean Mandarin (Fun,

2005), where the trisyllabic expletive *TaMaDe* (TMD) can always appear at the beginning and the end of Chinese utterances.

Secondly, when E is placed in monomorphemes or translated names with three or more syllables, the position before the last syllable is always prohibited from insertion. For instance, it is good to have *ping4* E *gwo2* ‘apple’, *so1* E *fu4 lei4* ‘soufflé’, *maa5* E *loi6 sai1 aa3* ‘Malaysia’, *maa5 loi6* E *sai1 aa3*, but never **so1 fu4* E *lei4* or **maa5 loi6 sai1* E *aa3*. It is interesting to us why this second last position of a string is always disallowed when the number of syllables existing in the string turn to three or above.

Thirdly, there is more than just the syntactic boundary that is in action to determine possible insertion place of E in strings with smaller internal structures. For example, *gam3 ci1 sin3* ‘so crazy’ (1+2) allows expletive insertion in both slots, resulting *gam3* E *ci1 sin3*, which is at the syntactic boundary, and *gam3 ci1* E *sin3*, where E breaks the adjective being modified by adverb *gam3* ‘so’. On the other hand, with *ci1 sin3 lou2* ‘crazy man’ (2+1), E is again allowed between ‘crazy’, resulting *ci1* E *sin3 lou2*, but not at the syntactic boundary **ci1 sin3* E *lou2*. It becomes intriguing to know why the insertion at the slot where the syntactic boundary cuts is not acceptable. These two strings with ‘crazy’ and the pair with ‘apple’ exemplified in the previous section show there is more to consider in this epenthesis puzzle than just syntax.

The central research question is therefore as given in (1).

(1) Central Research Question

What are the principles for expletive insertion in Hong Kong Cantonese?

1.3 Structure of the dissertation

This dissertation has a total of seven chapters to address the puzzle. The first chapter establishes the puzzle and the types of questions we need to provide answers. The second chapter will provide the social context of expletive epenthesis in Hong Kong Cantonese. The third chapter will show data of expletive epenthesis in the language, patterns and their respective categories. The fourth chapter is about some previous studies of curse words insertion done for other languages like American English (Hammond, 1999) and Singaporean Mandarin (Fun, 2005). The fifth chapter introduces the theoretical framework used in this dissertation, Optimality Theory (OT), and the relevant constraints that are in action to select the insertion place for expletive in our target language. The sixth chapter will talk about the relevant constraints in greater detail by providing more examples and establish their ranking based on established arguments. The last chapter will conclude the OT attempt and provide insights and implications for future research.

Cantonese syllables are rendered in Jyutping whenever possible, a romanisation system designed by the Linguistic Society of Hong Kong, to facilitate easy reading. Use of symbols and tableaux will be explained clearly when the illustrations come to sight.

Chapter Two

Background

Cantonese is not one language that is easy to learn given its complex tonal system and many allophones. Some say in order to master a language, the first component you need to learn to grasp the meaning is the swear words. Cantonese profanity and swear words (粗口 *cou1 hau2* ‘lit. thick mouth’) is quite unique in this sense. Its usage is very common among Cantonese speakers in Hong Kong, Macau and Chinatowns around the world. Cantonese speakers say profanity when they want to insult or abuse someone, intensify their meanings and express emotions.

There are mainly two types of Cantonese profanity, namely expletive utterances and expletive epenthesis. The first category are single utterances where they can be used separately once you know and learn the meaning of those phrases, for instance *diu2 nei5 lou5 mou2* ‘motherfucker’ and *ding2 nei5 go3 fai3* ‘damn you’. It is more simple when compared to the second category since to use the second category effectively, it requires more ‘sophistication’. The process of inserting a curse word or expletive into natural utterances to achieve the above purposes is defined as expletive epenthesis. You have to know the functions and properties of those monosyllabic words in order to be cursing in the second category fluently. *diu2*, *lan2*, *kau1* and *ceon1* meaning ‘male genitalia’ and *hai1* meaning ‘female genitalia’ are the five monosyllabic words that Hong Kong Cantonese speakers always insert in normal phrases or utterances. Some speakers may find these five words too strong and aggressive to them so they resort to euphemistic

options like English letters Q and L, but the mechanism of insertion remains the same and the word options can be used interchangeably.

Due to the pressure of the society and the conservative culture of the teaching environment in Hong Kong, expletive utterances or expletive epenthesis are never officially taught in a classroom setting. However, Cantonese language users, mostly adults, are able to do monosyllabic expletive epenthesis and have a consensus of whether the epenthesis is acceptable or not. It is a productive process and speakers can create new utterances with the expletive inserted once they learn the process. The purpose of this dissertation is to figure out what is regarded as natural, or ‘sound right’ for the second category of Cantonese profanity, namely expletive epenthesis in the Hong Kong variety of Cantonese and to propose the mental knowledge, in particular the prosodic structure of the specific variety, that lies behind to create such a phenomenon.

The following chapter will provide observations of the possible and impossible combinations of Hong Kong Cantonese strings with the monosyllabic expletive.

Chapter Three

Observations

The chapter states patterns of expletive epenthesis in different kinds of Hong Kong Cantonese strings, specifically monomorphemes and polymorphemes, according to the intuition of the author, who is a native speaker and have lived in Hong Kong for more than twenty four years using the language daily and frequently.

3.1 Monomorphemes

We shall first take a look at monomorphemes with different number of syllables and their expletive insertion patterns. Let's begin the account by studying the insertion of a monosyllabic expletive in disyllabic monomorphemes. Below include four disyllabic monomorphemes, being 'Hong Kong', 'apple', 'spider' and 'butterfly' which are all names or noun phrases.

(2) E in disyllabic monomorphemes

(a) 香港 hoeng1 gong2 'Hong Kong'	*E香港	香E港	*香港E
(b) 蘋果 ping4 gwo2 'apple'	*E蘋果	蘋E果	*蘋果E
(c) 蜘蛛 zi1 zyu1 'spider'	*E蜘蛛	蜘E蛛	*蜘蛛E
(d) 蝴蝶 wu4 dip6 'butterfly'	*E蝴蝶	蝴E蝶	*蝴蝶E

From (2), we can see that the monosyllable expletive cannot be put at the front or at the end of these disyllabic strings. For instance, prefixing or suffixing E to disyllabic name *hoeng1 gong2* 'Hong Kong' is ungrammatical. There has to a slot for insertion of the

expletive so the one and only possible place is then filled, so E has to go in between the two syllables. It is because native speakers have to insert the monosyllabic expletive to disyllabic strings and they need to resort to the less awkward option, that is in between the two syllables, to complete the epenthesis process. When there are more slots available, for instance in trisyllabic strings (3) below, there is a slight difference in pattern and the author prefer one slot within the syllables for insertion to the other.

(3) E in trisyllabic monomorphemes

(a) 梳乎厘 so1 fu4 lei4 ‘soufflé’	*E梳乎厘 梳E乎厘 *梳乎E厘 *梳乎厘E
(b) 士巴拿 si6 baa1 naa4 ‘spanner’	*E士巴拿 士E巴拿 *士巴E拿 *士巴拿E
(c) 朱古力 zyu1 gu1 lik1 ‘chocolate’	*E朱古力 朱E古力 *朱古E力 *朱古力E
(d) 新加坡 san1 gaa3 bo1 ‘Singapore’	*E新加坡 新E加坡 *新加E坡 *新加坡E

(3) shows that again E cannot be inserted at the beginning or at the end of these trisyllabic monomorphemes. The second last slot of the string is also not preferred (*so1 fu4 E lei4) and that E can only be put after the first syllable (so1 E fu4 E lei4). Since they are monomorphemes which must go together to give meanings, it is interesting to see how they all prefer the insertion places within the construction, never before or after it. And among the two available slots, we have to figure out why the slot after the first syllable sounds better than that after the second syllable.

The following illustrations (4) and (5) will leave out the insertion exemplifications with the beginning and the end options to save space since we already learn that it is

impossible to do so with disyllabic and trisyllabic monomorphemes. Learning that it is unacceptable to prefix or suffix an expletive, we will next look at monomorphemes with four or more syllables.

(4) E in quadrisyllabic monomorphemes

(a) 馬來西亞 maa5 loi6 sai1 aa3 'Malaysia'	馬E來西亞	馬來E西亞	*馬來西E亞
(b) 羅馬尼亞 lo4 E maa5 lei4 aa3 'Romania'	羅E馬尼亞	羅馬E尼亞	*羅馬尼E亞
(c) 斯洛伐克 si1 lok3 fat6 hak1 'Czechoslovakia'	斯E洛伐克	斯洛E伐克	*斯洛伐E克
(d) 委內瑞拉 wai2 noi6 sei6 laai1 'Venezuela'	委E內瑞拉	委內E瑞拉	*委內瑞E拉

Let's take (4a) 'Malaysia' for illustration. Ignoring the slots before and after the string for the moment, there is logically three slots for insertion of the expletive 'Ma- (i)_ -lay- (ii)_ -si- (iii)_ -a'. (i) and (ii) are acceptable for E while (iii) is not and sound pretty weird to the author. This preference is very interesting to me. For such a monomorpheme as simple as this, why would there be such a dislike to the (iii) slot. (5) with monomorphemes with five syllables shows a similar liking.

(5) E in monomorphemes with five syllables

(a) 柴可夫斯基 caai4 ho2 fu1 si1 gei1 'Tchaikovsky'	柴E可夫斯基 柴可夫E斯基	柴可E夫斯基 *柴可夫斯E基
(b) 斯洛文尼亞 si1 lok3 man1 nei4 aa3 'Slovenia'	斯E洛文尼亞 斯洛文E尼亞	斯洛E文尼亞 *斯洛文尼E亞

(c) 馬達加斯加 maa5 daat6 gaa1 si1 gaa1 'Madagascar'	馬E達加斯加 馬達加E斯加	馬達E加斯加 *馬達加斯E加
(b) 沙地阿拉伯 saa1 dei6 aa3 laai1 baak3 'Saudi Arabia'	沙E地阿拉伯 沙地阿E拉伯	沙地E阿拉伯 *沙地阿拉E伯

In (4) and (5), E can go in any slots within the monomorphemes with four or five syllables except before the last syllable of the constructions. It sounds very awkward to the author if the second last syllable of these combinations is the expletive, but it would not be the case if the expletive is in other slots. Among the acceptable combinations, the author thinks there are some that sound 'better' than others. This intuition of preference will be revisited and explained in chapter six when we talk about the constraints ranking. At this point, we can generalise the patterns of expletive epenthesis in monomorphemes to the following statements in (6).

(6) Patterns of expletive epenthesis in monomorphemes

(a) For disyllabic monomorphemes, E goes in between.

(b) For monomorphemes with three or more syllables, the only unacceptable insertion slot is before the last syllable.

Keeping these confusing puzzles, the most important point being the unacceptability of monosyllabic expletive insertion at the second last position of a monomorphemic string with three or more syllables, we have for monomorphemes, we will move on to polymorphemes and other constructions.

3.2 Polymorphemes

The insertion situation is different in polymorphemes than that of monomorphemes. We will start by looking at polymorphemes with two syllables of different word categories.

(7) E in disyllabic polymorphemes

(a)	(i) 痴線 ci1 sin3 ‘crazy’	*E痴線	痴E線	*痴線E
	(ii) 唔係 m4 hai6 ‘cannot’	*E唔係	唔E係	*唔係E
	(iii) 食飯 sik6 faan6 ‘eat rice’	*E食飯	食E飯	*食飯E
	(iv) 寫書 se2 syu1 ‘write book’	*E寫書	寫E書	*寫書E
(b)	(i) 人死 jan4 sei2 ‘people die’	*E人死	*人E死	*人死E
	(ii) 我係 ngo5 hai6 ‘I am’	*E我係	*我E係	*我係E
	(iii) 你食 nei5 sik6 ‘you eat’	*E你食	*你E食	*你食E
	(iv) 一二 jat1 ji6 ‘one two’	*E一二	*一E二	*一二E

The only possible insertion place for these polymorphemes in (7a) is similar to that of disyllabic monomorphemes, that is in between the two syllables. However, for strings in (7b), E is unacceptable for all three slots. This is because that (7a) are word categories or phrases, like adjectival phrases and verb phrases, while (7b) are not phrases, but clauses or sentences. It appears that expletive epenthesis is only possible within word categories and phrases. We should remember this and consider only data that are words or phrases with more syllables, as shown in the following.

(8) E in left-branching (2+1) trisyllabic polymorphemes

(a) 蘋果樹 ping4 gwo2 syu6 'apple tree'	*E蘋果樹 蘋E果樹 *蘋果E樹 *蘋果樹E
(b) 蘋果批 ping4 gwo2 pai1 'apple pie'	*E蘋果批 蘋E果批 *蘋果E批 *蘋果批E
(c) 蘿蔔糕 lo4 baak6 gou1 'turnip cake'	*E蘿蔔糕 蘿E蔔糕 *蘿蔔E糕 *蘿蔔糕E
(d) 痴線佬 ci1 sin3 lou2 'crazy man'	*E痴線佬 痴E線佬 *痴線E佬 *痴線佬E
(e) 賭輸馬 dou2 syu2 maa5 'lose in horse race bet'	*E賭輸馬 賭E輸馬 *賭輸E馬 *賭輸馬E
(f) 打死人 daa2 sei2 jan4 'kill a person by hitting'	*E打死人 打E死人 *打死E人 *打死人E

In (8) where strings are in 2+1 composition, E is allowed to break the first two syllables but is prohibited in the space where the syntactic boundary cuts. The syntactic structure is not at work for these constructions to find acceptable insertion places so it is one of the reasons why we need to appeal to other forces.

(9) E in right-branching (1+2) trisyllabic polymorphemes

(a) 食蘋果 sik6 ping4 gwo2 'eat apple'	*E食蘋果 食E蘋果 食蘋E果 *食蘋果E
(b) 紅蘋果 hung4 ping4 gwo2 'red apple'	*E紅蘋果 紅E蘋果 紅蘋E果 *紅蘋果E
(c) 大西瓜 daai6 sai1 gwaa1 'big watermelon'	*E大西瓜 大E西瓜 大西E瓜 *大西瓜E
(d) 做功課 zou6 gung1 fo3 'do homework'	*E做功課 做E功課 做功E課 *做功課E

(e) 禁痴線 gam3 ci1 sin3 'so crazy'	*E禁痴線 禁E痴線 禁痴E線 *禁痴線E
(f) 唔可以 m4 ho2 ji5 'not okay'	*E唔可以 唔E可以 唔可E以 *唔可以E

Trisyllabic polymorphemes in (9) is in a structure of 1+2 and E can be inserted in the syntactic boundary, contrary to the pattern shown in (8). E again breaks the last two syllables which belong together like those first two syllables in (8); data (8) and (9) seem to be contradictory in this sense.

In (7) to (9), E is not allowed at the beginning or the end of these polymorphemes. Using a similar presentation with monomorphemes for easy reading, the following data on polymorphemes with more syllables will leave out the options of the front and the back as well.

(10) E in quadrisyllabic polymorphemes

(a)	right-branching (1+3)	
(i)	整蘋果批 zing2 ping4 gwo2 pai1 'make apple pie'	整E蘋果批 整蘋E果批 *整蘋果E批
(ii)	食朱古力 sik6 zyu1 gu1 lik1 'eat chocolate'	食E朱古力 食朱E古力 *食朱古E力
(b)	(2+2)	
(i)	服務態度 fuk6 mou6 taai3 dou6 'service attitude'	服E務態度 服務E態度 服務態E度
(ii)	新鮮蘋果 san1 sin1 ping4 gwo2 'fresh apple'	新E鮮蘋果 新鮮E蘋果 新鮮蘋E果

(iii) 超甜蘋果 ciu1 tim4 ping4 gwo2 'super sweet apple'	超E甜蘋果 超甜E蘋果 超甜蘋E果
(iv) 種植蘋果 zung3 zik6 ping4 gwo2 'plant apple'	種E植蘋果 種植E蘋果 種植蘋E果
(c) left-branching (3+1)	
(i) 朱古力餅 zyu1 gu1 lik1 beng2 'chocolate cookie'	朱E古力餅 *朱古E力餅 *朱古力E餅
(ii) 菲律賓人 fei1 leot6 ban1 jan4 'Filipino'	菲E律賓人 *菲律E賓人 *菲律賓E人

There are three possible internal structure for quadrisyllabic strings, being right-branching, 2+2 and left-branching. The three categories exhibits different patterns of expletive insertion as exemplified in (10). For 1+3 polymorphemes in (10a), the insertion is unacceptable before the last syllable. All three slots can be filled with the monosyllabic expletive for the 2+2 constructions in (10b). Left-branching trisyllabic polymorphemes only allow the insertion of E to be after the first syllable.

As demonstrated in tables (2) to (10), we can have the following generalisation in (11):

(11) Overall observations of expletive epenthesis in Hong Kong Cantonese

(a) Expletives cannot come before and after a string of any kind.

(b) Expletive epenthesis is only allowed within phrases.

(c) For monomorphemes, when they have three or more syllables, the slot before the last syllable is not acceptable. All other slots are acceptable.

(d) For polymorphemes, syntax and more forces are at work to determine possibilities.

Chapter Four

Previous studies

Expletive insertion is necessarily prompted by speakers' intuition of the language. Other than that in Hong Kong Cantonese as rendered in the previous chapter, expletive epenthesis also happens in stress-timed languages like English.

Hammond (1999) does a research on the infixation of expletive *fuckin'* in American English words. He states that discussing offensive speech may be offending to some readers but there are compelling reasons to study this language phenomenon. The informal and slangy registers show a sensitivity to the prosodic structure of a language and that could not be a consequence of explicit instruction in which makes the phenomenon remarkable. It is important to know that informal speech like curse words and expletives are no less capable of showing important linguistics facts.

Hammond (1999) considers disyllabic words with the possible expletive infixation and finds out that the infixation is only possible with words that has the secondary stress on the first syllable and the primary stress on the second syllable, that is the primary stress must follow the infixated expletive. He then demonstrates with more examples that has three syllables and generalises that firstly, expletive infixation can happen only when the primary stress occurs to the right of the infixation; secondly, infixation occurs when there is a foot on either side around the infixation and thirdly, infixation cannot interrupt a root. The three rules for expletive infixation in American English can be easily understood with the word *influenza* ([,ɪn][flu]['enzə] where square brackets indicate the

foot boundary). *influenza* has its primary stress on *-en-* and secondary stress is *in-*. Inserting *fuckin'* in between *in-* and *-flu-* is acceptable because the primary stress follows although not immediately after, there is a foot *-in-* and two feet *-flu-* and *-enza-* around *fuckin'* and it does not break down any of the three feet. *influ-fuckin'-enza* will also be okay but for *influen-fuckin'-za*, it is not acceptable because the expletive is now on the right side on the syllable carrying the primary stress and it breaks the third foot which has two syllables. Without positioning the foot as part of the way people organise prosody unconsciously, it would be rather difficult to capture these patterns and even come up with rules to figure out how it works.

The above three rules focus on the functions of stress and foot structure on expletive infixation in American English. Expletive Epenthesis in Hong Kong Cantonese is similar to this case in American English in a sense that they are both infixations. Doing an analysis on the Hong Kong Cantonese data can help to provide evidence on what the foot structure of the language is like and what kinds of stress or rhythm the Hong Kong Cantonese language users prefer. Since this project on Hong Kong Cantonese aims to predict the infixation cases with expletive of different number of syllables, it would have insights on expletive epenthesis in other languages like English since the existing analysis by Hammond (1999) only deals with one expletive *fuckin'* which has two syllables. Though studies have started to show that syllable-timed language like Chinese do also have stress (Hsieh, 2020), it is unconventional to say Cantonese or Mandarin have stress. This is where the foot structure becomes so prevalent to talk about both stress-timed and syllable-timed languages.

Fun performs an interesting analysis in 2005 to explain how the expletive *TaMade* (TMD) is inserted in Singaporean Mandarin. Examining strings with two to five syllables, it is found that for constructions that have the same internal syntactic structure, it is possible for them to have different places of expletive insertion, and vice versa and this is due to the very complex interactions of syntax and foot. Fun (2005) provides an Optimality Theory account which makes use of six ranked constraints (shown in (12)), aiming to explain how Singaporean Mandarin speakers insert an expletive to a monomorpheme only when it has at least four syllables and the insertion of expletive for both monomorphemes and polymorphemes is always possible at the boundary of words and phrases. The constraints is about the preference of certain foot and how the boundary of phrases, foot and syntactic structures align with one another.

(12) Constraints (Fun, 2005)

(a) **DF** (Disyllabic foot)

It is preferred to have two syllables in one foot.

(b) **FF** (Final foot)

The final foot cannot be monosyllabic.

(c) **ALIGN-PHRASE** (Phrase alignment)

Any phrase boundary has to be a foot boundary.

(d) **ALIGN-SYNTAX** (Highest syntactic boundary alignment)

The boundary of the highest level of a syntactic representation has to be a foot boundary.

(e) **ALIGN-WORD** (Word alignment)

Any word boundary has to be a foot boundary.

(f) **ALIGN-FOOT** (Foot alignment)

Any foot boundary has to be a word boundary.

These six constraints have been used to choose the optimal prosodic structures with regard to expletive insertion and they have their effects in choosing the preferred foot and its alignment of boundary cut with the syntactic structures. The ranking is settled to be **ALIGN-PHRASE** >> **FF** >> **ALIGN-SYNTAX**; **DF** >> **ALIGN-FOOT**; **ALIGN-WORD** with careful data examination and exemplification by Fun (2005).

An OT analysis can help to explain why certain places for expletive insertion is not acceptable and also predict all the possible combinations for different kinds of phrases effectively once we settle with a factual rank of constraints. This set of constraints by Fun (2005) is very useful for reference and can help us to find a solution for the Hong Kong Cantonese data but is inadequate since the analysis on Singaporean Mandarin is primarily based on the fact that the expletive TMD can be inserted at the boundary of any constructions. Taking an example from Fun (2005), the monomorphemic quadrisyllabic string *LuoMaNiYa* ‘Romania’ can have three possible insertion, being *TMDLuoMaNiYa*, *LuoMaNiYaTMD* and *LuoMaTMDNiYa*, showing that the insertion is possible at the beginning and ending boundary (effects of the alignment constraints) and that it is preferred to have two syllables in a foot (effects of DF and FF). It is not the case for Hong Kong Cantonese where its expletive is kind of like an infix as shown in

chapter three and the insertion place needs not be at the boundary of any word and phrase; it is also yet unclear whether Hong Kong Cantonese also prefers a disyllabic foot. Other constraints have to be introduced to put the relationship and meanings of the parts of the constructions into consideration so as to better describe and analyse the Hong Kong Cantonese data.

Chapter Five

Theoretical framework

5.1 Optimality Theory

Optimality Theory (OT) is introduced by Alan Prince and Paul Smolensky in the 1990s. It is a theory about the organic combinations and interactions of the specification and universality of languages, in which it becomes one of the most important and influential linguistics theories today. Though originally developed from generative phonology, the principles of OT have also been applied to studies of syntax, morphology, pragmatics, language change, and other areas throughout the years. In the theory, surface forms of language reflect the resolution of conflicts between competing and ranked constraints. OT holds that all languages have a set of constraints which produce the basic phonological and grammatical patterns. The constraints are universal and innate while the ranking of the constraints have to be acquired, helping to explain the varieties of many languages.

Optimality Theory can help to provide a powerful argument to the explanation and solution of the expletive epenthesis puzzle in Hong Kong Cantonese. (13) is the framework of Optimality Theory to explain how the grammar works.

(13) Optimality Theory framework

Input → Generator → Candidates → Evaluator → Output

The function Generator is a crucial part of the grammar; it produces a set of candidates consistent with a given input from the lexicon that contains all the underlying forms. The produced candidates then enter the Evaluator in which the Evaluator system contains with a system of ranked constraints. When accessing the candidates, the system will evaluate the candidates in terms of their relative harmony, or degree of success and more and more candidates will be eliminated according to the constraint preference until the maximally harmonic candidate is selected. This output is optimal and the rest candidates are discarded. This assessment can be presented in a tableau for easy analysis and reading, as shown in (14) below. The upper left corner shows inputs, while the candidates produced by the Generator will listed on the first column from up to down indexed. The constraints are listed in terms of their ranking from left to right.

(14) Optimality Theory assessment tableau

Inputs	Constraint 1	Constraint 2	Constraint 3	Constraint 4
i. candidate A	*!		*	
ii. candidate B		*!		
iii. candidate C			**!	
iv. candidate D			*	*!
v. candidate E			*	

The asterisks (*) in (14) indicates the number of violations of the candidate to a given constraint while the exclamation marks indicates that the violation is critical so the specific candidate is eliminated. As mentioned, the constraints from higher to lower rank are listed in the column from left to right. Candidate A violates the highest-ranking

constraint so it is the first to be eliminated in the assessment process. Since candidate A is already eliminated by the effect of constraint 1, its violations to other constraints will not affect the fact that it will never be chosen as the optimal output. The cells that are not relevant to the selection process anymore are filled with colour grey. Candidate B violates the second constraint so is then eliminated after candidate A. Candidates C, D and E all violate constraint 3 but only the violation of candidate C is critical since it violates constraint 3 for two times. Both candidates D and E violate constraint 3 once, however, candidate D violates the final constraint four, it is then crossed out. Although candidate E violates constraint 3, it is among the five candidates with the least violations, it will be selected as the optimal output and be pointed with a little hand symbol.

Constraints have a universal tendency which can explain the similarities of different languages around the world, while it is the different ranking of the constraints show how different languages appear differently to particular linguistics characteristics. The following section will go into talking about types of constraints we need to use for the solution of the Hong Kong Cantonese expletive epenthesis puzzles.

5.2 Relevant constraints

Given the observations and patterns as listed in chapter three, the following five constraints in (15), (16) and (18) are used to address the puzzle of expletive epenthesis in Hong Kong Cantonese.

(15) Faithfulness constraint

(a) **MAX-IO** (Maximal input and output)

Assign one violation mark if an input does not have an output correspondent.

This constraint prohibits deletion and addition to the candidate other than the two inputs, being the monosyllabic expletive and the string in Hong Kong Cantonese that has to be infixes.

(16) Constraints on foot structure

(a) **PARSE-SYLL** (Parse syllables)

All syllables are parsed by foot.

(b) **FT BIN** (Foot binary)

Feet must be binary.

(c) **FF** (Final foot)

The final foot cannot be monosyllabic.

These constraints about the foot structure suggest the importance of parsing syllables to a foot in Hong Kong Cantonese and the kind of foot structure our target language Hong Kong Cantonese may mostly prefer. In Indonesian (McCarthy and Cohn, 1998), **PARSE-SYLL** and **FT BIN** are also active in determining the stress and foot structure in odd-parity words as shown in (17).

(17) FT BIN >> PARSE-SYLL, from /bicara/ ‘speak’

Candidates	FT BIN	PARSE-SYLL
i. \blacksquare bi [cára]		*
ii. [bì] [cára]	*!	

The surface form for the underlying representation /bicara/ is bi [cára] where the square brackets indicate the foot boundary. The constraint for preference of a binary foot is ranked higher than the need for parsing every syllable to a foot so candidate (17i) is selected even though the first syllable of it does not contribute to form a foot; this conflict between FT BIN and PARSE-SYLL is resolved by the ranking. A very different choice of optimal candidate will appear if the ranking is changed and there will be also other consequences. Different languages need different constraints and their related ranking to appear as how they are now on surface. The following is the last active constraint for this puzzle; it is the alignment constraint in (18).

(18) Alignment constraint

(a) **ALIGN EDGE [XP, X; PHP]**

For any XP and X, assign one violation mark for each syllable that separates the edgemoſt XP or X ſyllable with the correſponding edge of the phonological phrase or word.

The Generalised Alignment definition (19) of alignment constraints has played a key role in Optimality Theoretic approaches (Hyde, 2012). It is about the edge of every

category coinciding with the edge of another category. A violation mark will be given for every intervening separator between edges which prevent coincidence of edges of two categories and cause misalignment.

(19) Generalised Alignment

ALIGN (ACat1, Edge1, ACat2, Edge2, SCat)

The Edge1 of every ACat1 coincides with the Edge2 of some ACat2. Assess a violation mark for every SCat that intervenes between edges that fail to coincide.

All of these constraints exist in the Universal Grammar and are used in Optimality Theory for analysing other languages. The following chapter will talk about each constraint and their power in details and explain how they function in our puzzle. Example data will be listed to propose a suitable ranking of the active constraints.

Chapter Six

OT analysis and constraints ranking

This chapter starts by analysing monomorphemes to introduce the constraints mentioned in the previous section. Before going into the data, two constraints MAX-IO and PARSE-SYLL are proposed to be ranked the highest together. The faithfulness constraint MAX-IO must rank very high because we are to consider the expletive infixation process to a given Hong Kong Cantonese string. We will not consider candidates that occur deletion or addition of syllables to them.

The concept of foot is critical to explaining the distribution of the monosyllabic expletive in Hong Kong Cantonese, so we know that every syllable needs to be parsed to a foot and that PARSE-SYLL is ranked very high as well. Every syllables in a candidate will be parsed to a foot and we will not consider candidates which have syllables that remain unfooted. We should remember that these two constraints are visibly active in Hong Kong Cantonese but we will not include them in the following demonstration because they are clearly ranked high above the others. Keeping this in mind, let's recall from chapter three the expletive epenthesis patterns to disyllabic monomorphemes.

6.1 Monomorphemes

For instance *ping4 gwo2* 'apple, when the expletive is inserted, it has to only be in the middle of two syllables, resulting *ping4 E gwo2*; prefixing or suffixing the expletive to the disyllabic string is not permitted.

(20) Optimal candidate for E + *ping4 gwo2* ‘apple’

E + (NP ping4 gwo2)	ALIGN EDGE
i. (NP E [ping4 gwo2])	*!
▣ ii. (NP [ping4] E [gwo2])	
iii. (NP [ping4 gwo2] E)	*!

Square brackets indicate the foot boundary while round brackets indicate the boundary of the phonological phrase while the syllables that form a foot stay inside the brackets. The three candidates in (20) are the three possibilities that we can insert a monosyllabic expletive (E) in the disyllabic noun phrase *ping4 gwo2*. Both inserting E before the noun phrase (20i) and attaching after the noun phrase (20iii) violates ALIGN EDGE once respectively, making (20ii) the best candidate. This is a quite simple and small example to start with so we need to consider phrases with more syllables next to put other constraints in play.

Here we have *si6 baa1 naa4* ‘spanner’ where the only possible insertion place for the monosyllabic expletive is after the first syllable (*si6 E baa1 naa4*).

(21) Optimal candidate for E + *si6 baa1 naa4* ‘spanner’

E + (NP si6 baa1 naa4)	ALIGN EDGE	FF
i. (NP E [si6 baa1 naa4])	*!	
▣ ii. (NP [si6] E [baa1 naa4])		

iii. (NP [si6 baa1] E [naa4])		*!
iv. (NP [si6 baa1 naa4] E)	*!	

(21) selects (ii) to be the best candidate. After eliminating the prefixed and suffixed candidates using the same constraint ALIGN EDGE like (20). FF is in action to compete candidates (21ii) and (21iii). The final foot must not be binary so (iii) violates FF critically. We also cannot differentiate whether ALIGN EDGE and FF rank higher than the other because the optimal candidate (21ii) do not have any violations. We shall move on to consider a quadrisyllabic monomorpheme to see if this piece of data helps us to make a differentiation of these two discussed constraints.

For a monomorpheme with four syllables like *maa5 loi6 sai1 aa3* ‘Malaysia’, we can imagine five available slots for expletive insertion, that is: (22i) E *maa5 loi6 sai1 aa3*, (22ii) *maa5* E *loi6 sai1 aa3*, (22iii) *maa5 loi6* E *sai1 aa3*, (22iv) *maa5 loi6 sai1* E *aa3* and (22v) *maa5 loi6 sai1 aa3* E.

(22) Optimal candidate for E + *maa5 loi6 sai1 aa3* ‘Malaysia’

E + (NP <i>maa5 loi6 sai1 aa3</i>)	ALIGN EDGE	FF	FT BIN
i. (NP E [<i>maa5 loi6 sai1 aa3</i>])	*!		*
ii. (NP [<i>maa5</i>] E [<i>loi6 sai1 aa3</i>])			**!
iii. (NP [<i>maa5 loi6</i>] E [<i>sai1 aa3</i>])			
iv. (NP [<i>maa5 loi6 sai1</i>] E [<i>aa3</i>])		*!	**
v. (NP [<i>maa5 loi6 sai1 aa3</i>] E)	*!		*

(22i) and (22v) violate the high ranking ALIGN EDGE and are eliminated at first again. The last foot must not be monosyllabic and further this constraint further cross out candidate (22iv). An additional constraint introduced here in this tableau is about keeping each foot to be binary so with a foot of one syllable and the other foot of three syllables, (22ii) is eliminated here to find the optimal candidate (22iii) (_{NP} [maa5 loi6] E [sai1 aa3]).

From the observation in (4a), we learn that both *maa5 E loi6 sai1 aa3* and *maa5 loi6 E sai1 aa3* are acceptable. Here we need to bring back one pattern that to the author, there are some combinations of the inputted string and the monosyllabic expletive sounding better than the others. In this case for ‘Malaysia’, inserting E in the middle to form two separate foot containing two syllables each (*maa5 loi6 E sai1 aa3*) is the one that to the author is more natural than putting the expletive after the first syllable to result in *maa5 E loi6 sai1 aa3*. This makes the constraint FT BIN ranking lower than the other constraints as listed because even a candidate violates it, it could still be an attested surface form in the natural language world.

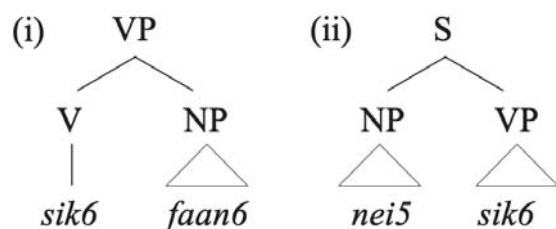
These three examples of monomorphemes in (20-22) demonstrate how ALIGN EDGE, FF and FT BIN work in Hong Kong Cantonese. The fact that the author prefers some attested forms than the others can also partially be explained in the OT analysis like (22). This preference shows a little sense of how the introduced constraints rank in the

language. We shall move on to polymorphemes for further demonstrations and an attempt to explain the complex interaction of foot structure and syntax.

6.2 Polymorphemes

This section discusses the insertion of the monosyllabic expletive to polymorphemes of different number of syllables and explain why some insertions are acceptable by looking at the optimal candidate produced from the Optimality Theory analysis. Polymorphemes are different and more complicated for analysis in a sense that the analysis must consider both the phonological phrase and syntactic phrase and their interactions with the foot structure. In (7aiii), *sik6 faan6* ‘eat rice’ can be inserted with the expletive to become *sik6 E faan6*, but it is not the case for (7biii), *nei5 E sik6* ‘you eat’ is not acceptable. This is because *sik6 faan6* and *nei5 sik6* have different syntactic trees as shown below in (23):

(23) Tree diagrams for *sik6 faan6* ‘eat rice’ and *nei5 sik6* ‘you eat’



From tree diagrams in (23), we see that the two disyllabic polymorphemes have different syntactic structures while expletive insertion can only happen in phrases like (23i) but not clauses or sentences like (23ii). This is also why the alignment constraints

specified for use here in the puzzle mention the area for application is only in phonological phrases. (24) shows the tableau for selecting the candidate for *sik6 faan6* since the constraints can apply to its case.

(24) Optimal candidate for E + *sik6 faan6* ‘eat rice’

E + (VP v sik6 (NP faan6))	ALIGN EDGE	FF	FT BIN
i. (VP E [v sik6 (NP faan6)])	*!		*
ii. (VP [v sik6] E [(NP faan6)])		*	**
iii. (VP [v sik6 (NP faan6)] E)	*!		*

Candidate (24ii) can be easily selected due to the prominence of ALIGN EDGE. Although this candidate violate more times on FF and than FT BIN (24i) and (24iii), those violations are not relevant anymore since ALIGN EDGE is ranked high. This is one verb phrase example in (24) while (25) will show that AdjP like *ci1 sin3* ‘crazy’ will do the similarly under the same mechanism.

(25) Optimal candidate for E + *ci1 sin3* ‘crazy’

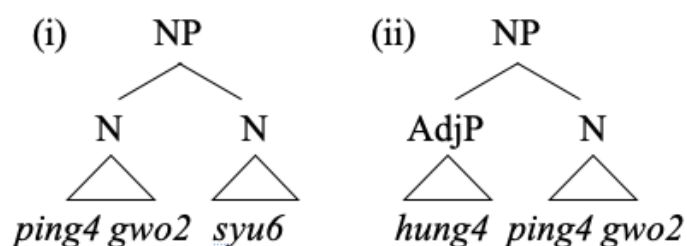
E + (AdjP (VP ci1) (NP sin3))	ALIGN EDGE	FF	FT BIN
i. (AdjP E [(VP ci1) (NP sin3)])	*!		*
ii. (AdjP [(VP ci1)] E [(NP sin3)])		*	**
iii. (AdjP [(VP ci1) (NP sin3)] E)	*!		*

ci1 sin3 is proposed to be composed of the verb phrase *ci1* ‘cross’ and noun phrase *sin3* ‘wires’. (25i) and (25iii) are eliminated by the powerful ALIGN EDGE since the very strong dislike towards prefixing and suffixing the expletive to any kinds of strings. The violations in grey cells are not prevalent to the choosing of optimal and attested form candidate (25ii). From all the tableaux we have, we know that ALIGN EDGE is the most important constraint among the three. Constraint mentioning the final foot (FF) should not be monosyllabic follows ALIGN EDGE while constraint about binary foot (FT BIN) is listed last due to the fact that even violations occur, some candidates may still be attested. The current constraint ranking we have is put in (26).

(26) ALIGN EDGE >> FF >> FT BIN

We shall move on to more complex polymorphemes. Recall from (8) and (9) rendering the expletive epenthesis pattern on trisyllabic polymorphemes, we will now make use of two trisyllabic strings involving *ping4 gwo2* ‘apple’ to compare and explain the difference insertion patterns. (27) shows the syntactic structures of the two strings, *ping4 gwo2 syu6* ‘apple tree’ and *hung4 ping4 gwo2* ‘red apple’.

(27) Tree diagrams for *ping4 gwo2 syu6* and *hung4 ping4 gwo2*



(27i) is a compound noun phrase that composes of two nouns, ‘apple’ and ‘tree’ while (27ii) is a noun phrase that has an adjectival phrase ‘red’ and the head noun ‘apple’ being modified. 2+1 structures like (27i) only allow $((\sigma_1 E \sigma_2) \sigma_3)$ but not $((\sigma_1 \sigma_2) E \sigma_3)$ while 1+2 structures like (27ii) accept both $(\sigma_1 E (\sigma_2 \sigma_3))$, but not $(\sigma_1 (\sigma_2 E \sigma_3))$. The insertion pattern for 2+1 structures can be explained using the existing three constraints up to this point of discussion, as exemplified in (28).

(28) Optimal candidate for E + ping4 gwo2 syu6 ‘apple tree’

E + (NP (N ping4 gwo2) N syu6)	ALIGN EDGE	FF	FT BIN
i. (NP E [(N ping4 gwo2) N syu6])	*!		*
ii. (NP [(N ping4) E [gwo2) N syu6])			*
iii. (NP [(N ping4 gwo2)] E [N syu6])		*!	*
iv. (NP [(N ping4 gwo2) N syu6] E)	*!		*

E attaching to the beginning (28i) and the end (28ii) separates the phonological phrase indicated by the round bracket and the foot indicated by the square brackets, making critical violations and therefore are eliminated at ALIGN EDGE. (28iii) has the last syllable ‘tree’ standing alone in the foot so FF is violated. The attested form *ping4 E gwo2 syu6* in candidate (28ii) is successfully selected. Things get a little trickier for 1+2 structures below. Both *hung4 E ping4 gwo2* and *hung4 ping4 E gwo2* are accepted and sound equally natural to the author. Why? (29) shows how (29ii) is selected as the most harmonic candidate as usual, but what about (29iii), which is also an attested form.

(29) Optimal candidate for E + hung4 ping4 gwo2 ‘red apple’

E + (NP (AdjP hung4) (N ping4 gwo2))	ALIGN EDGE	FF	FT BIN
i. (NP E [(AdjP hung4) (N ping4 gwo2)])	*!		*
ii. (NP [(AdjP hung4)] E [(N ping4 gwo2)])			*
iii. (NP [(AdjP hung4) (N ping4)] E [gwo2])		*!	*
iv. (NP [(AdjP hung4) (N ping4 gwo2)] E)	*!		*

One possible explanation is that given an internal structure like 1+2, expletive insertion is carried out again in the second and third syllable, so things will be like what happen to disyllabic strings. Since adding at the beginning or at the end is not acceptable, the expletive has to go in between, resulting a kind of conflicting pattern that the OT tableau cannot manage to explain yet. Until other constraints are found to be useful in explaining this situation and may need to be ranked on the same level as FF to make the number of violations the same so as to allow two attested forms, this problem needs to be noted and addressed in the future.

After looking through monomorphemes and polymorphemes of relatively few syllables, we are about all set to have our ranking of five constraints in the attempt to explaining puzzles in expletive epenthesis in Hong Kong Cantonese. We will verify the ranking with monomorphemes and polymorphemes of four or more syllables in the upcoming tableaux.

(30) Optimal candidate for E + caai4 ho2 fu1 si1 gei1 ‘Tchaikovsky’

E + (NP caai4 ho2 fu1 si1 gei1)	ALIGN EDGE	FF	FT BIN
i. (NP E [caai4 ho2 fu1 si1 gei1])	*!		*
ii. (NP [caai4] E [ho2 fu1 si1 gei1])			**
iii. (NP [caai4 ho2] E [fu1 si1 gei1])			*
iv. (NP [caai4 ho2 fu1] E [si1 gei1])			*
v. (NP [caai4 ho2 fu1 si1] E [gei1])		*!	
i. (NP [caai4 ho2 fu1 si1 gei1] E)	*!		*

Here is a translated name Tchaikovsky consisted of five Hong Kong Cantonese syllables presented in (30). (30ii), (30iii) and (30iv) are our attested forms if we recall from (5) chapter three and they are successfully selected as optimal candidates with the effects of ALIGN EDGE and FF. Two violations of FT BIN from (30ii) explains why (30iii) and (30iv) with one violation respectively sound better than it. This set of constraints, in particular constraint FF, comes straight on point in figuring out why the second last spot of any strings is prohibited from inserting the monosyllabic expletive. Let’s move on to polymorphemes with four syllables for further evidences.

(31) Optimal candidate for E + zing2 ping4 gwo2 pai1 ‘make apple pie’

E + (VP (V zing2) (NP (N ping4 gwo2) N pai1))	ALIGN EDGE	FF	FT BIN
i. (VP E [(V zing2) (NP (N ping4 gwo2) N pai1)])	*!		*
ii. (VP [(V zing2)] E [(NP (N ping4 gwo2) N pai1)])			*
iii. (VP [(V zing2) (NP (N ping4) E [gwo2] N pai1)])			*
iv. (VP E [(V zing2) (NP (N ping4 gwo2)] E [N pai1]))		*!	*
v. (VP [(V zing2) (NP (N ping4 gwo2) N pai1)] E)	*!		*

The first and fifth candidates with the prefixed and suffixed expletive are again easily discarded with the high-ranked ALIGN EDGE. The second last spot for the expletive is not preferred so (31iv) is eliminated with the activity of FF. The two optimal candidates (31ii) and (31iii) selected are the attested forms we have. The next tableau show a 2+2 string which have all slots within the four syllables available for expletive epenthesis.

(32) Optimal candidate for E + san1 sin1 ping4 gwo2 ‘fresh apple’

E + (NP _(AdjP san1 sin1) (NP ping4 gwo2))	ALIGN EDGE	FF	FT BIN
i. (NP E [(AdjP san1 sin1) (NP ping4 gwo2)])	*!		*
ii. (NP _(AdjP [san1] E [sin1]) [(NP ping4 gwo2)])		*!	**
iii. (NP [(AdjP san1 sin1)] E [(NP ping4 gwo2)])			
iv. (NP [(AdjP san1 sin1)] (NP [ping4] E [gwo2]))		*!	**
v. (NP [(AdjP san1 sin1) (NP ping4 gwo2)] E)	*!		*

It is worthwhile to talk about this 2+2 structure since its treatment would be like dealing with two disyllabic monomorphemes, which makes the only spot within the two syllables and the middle place of the whole polymorpheme are all available for expletive insertion. The tableau (32) indicates the optimal form is (32iii) where the expletive is inserted in the middle with two syllables on either side. This explains why the other attested forms (32ii) and (32iv) do not sound as ‘good’ as (32iii). (32iii) do not have any violations for all constraints. The next example in (33) is the last composition a quadrisyllabic polymorpheme can give, that is the 3+1 structure.

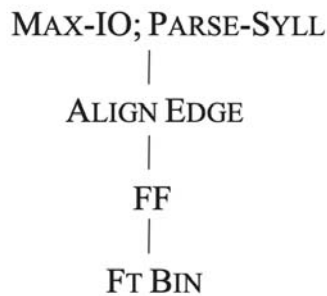
(33) Optimal candidate for E + *fei1 leot6 ban1 jan4* ‘Filipino’

E + (NP (N <i>fei1 leot6 ban1</i>) (N <i>jan4</i>))	ALIGN EDGE	FF	FT BIN
i. (NP E [(N <i>fei1 leot6 ban1</i>) (N <i>jan4</i>)])	*!		*
ii. (NP (N [<i>fei1</i>] E [<i>leot6 ban1</i>]) [(N <i>jan4</i>)])		*	**
iii. (NP (N [<i>fei1 leot6</i>] E [<i>ban1</i>]) [(N <i>jan4</i>)])		**!	**
iv. (NP [(N <i>fei1 leot6 ban1</i>)] E [(N <i>jan4</i>)])		*	**
v. (NP [(N <i>fei1 leot6 ban1</i>) (N <i>jan4</i>)] E)	*!		*

(33) includes a 3+1 example *fei1 leot6 ban1 jan4* ‘Filipino’ and the only attested form is *fei1 E leot6 ban1 jan4*. Prefixed and suffixed versions are discarded on the first constraint ALIGN EDGE without doubt. (33iii) violates FF for two times having *ban1* and *jan4* as monosyllabic final foot in the 3+1 structure, making its violations critical and being eliminated in this second stage. There is a challenge indicating by the bomb symbol in the tableau choosing the attested form (33ii) from (33ii) and (33iv) since they both violates FF once and the same number of the violations in FT BIN. One possible explanation to the difference of these two candidates that have not been successfully captured by the current model of constraints is that on the bigger noun phrase of four syllables, inserting the expletive before the last syllable *jan4* in (33iv) violates FF again in a larger sense or higher level.

With the above illustrations in chapter six, we have arrive at the suitable ranking for the five listed constraints. (34) presents the Hesse Diagram for the ranking.

(34) Hesse Diagram of the five constraints



As we have mentioned at the beginning of the chapter, MAX-IO and PARSE-SYLL are ranked high together since we are to consider expletive epenthesis to a given input and the concept of the foot structure is crucial to understanding the puzzle so every syllables among the inputs have to be parsed as members of a foot. ALIGN EDGE is what follows them since unlike what has been found by Fun (2005) in Singaporean Mandarin, Hong Kong Cantonese does not prefer expletive epenthesis at the beginning nor at the end of any kinds of strings. ALIGN EDGE stops the process from choosing expletive prefixed or suffixed candidates as the expletive separates the edgemost syllable with the corresponding edge of the phonological phrase or word. Given one special characteristic in the Hong Kong Cantonese expletive epenthesis puzzle is that the position before the last syllable of any strings is not acceptable for insertion, FF is in action to eliminate those combination which has their final foot as monosyllabic. FT BIN helps to understanding why some candidates, though attested, sound less natural than the other, which occurs the least violations to FT BIN. MAX-IO; PARSE-SYLL >> ALIGN EDGE >> FF >> FT BIN is our final constraint ranking to settle.

Chapter Seven

Conclusion

In this dissertation, the author has presented an account of the prosodic and foot structures in Hong Kong Cantonese and the ways it surfaces in the puzzles of expletive epenthesis in the language. The analysis is set within Optimality Theory, and the properties of constraints on faithfulness, foot structure and alignment are particularly relevant to our attempt.

There are residual issues that we still need to work on in the future. One of them is to find ways to present the preference of certain insertion slots to others and the Optimality Theory framework should be able to capture the degree of acceptability when more constraints are introduced to the ranking for Hong Kong Cantonese. Another issue is that data could be collected from more native speakers of different ages and groups to capture their intuition to the expletive epenthesis patterns. Although the author's impression could serve as reference, it is important to know that we have to include more speakers in the data to come up with more accurate generalisations. Since this is a vernacular, speakers' intuition could really vary a lot so the differences among speakers have to be taken into account. In the future, we could work with more strings other than the current set of monomorphemes and polymorphemes and can even have fun with sentences and clauses when we even react with more complex syntactic structures to see if our determined ranking still works.

Hong Kong Cantonese is probably one of the most ideal languages to study prosodic patterns since it is a colloquial language that reflects a lot of language knowledge and intuition native speakers have. This dissertation hopes to provide insights to the understanding of a very colloquial language phenomenon and a very elementary attempt of Optimality Theory analysis.

REFERENCES

- Fun, Cedric How Thern (2005). *Jiekou: xiandai hanyude liyan chuancha* [接口:现代汉语的詈言穿插] (*Interface: Expletive Insertion in Modern Chinese*). Graduation thesis, National University of Singapore.
- Hammond, Michael (1999). *The Phonology of English: an Optimality Theoretical Approach*. Oxford and New York: Oxford University Press.
- Hsieh, Feng-fan (2020). *Lun hanyu you cizhongyin* [论汉语有词重音] (*On the existence of word stress in Chinese*). Retrieved September 1, 2020 from <https://www.researchgate.net/publication/342918760>.
- Hyde, Brett (2012). Alignment constraints. *Natural Language and Linguistic Theory*. 789-836. Retrieved November 1, 2020 from onAcademic.com.
- McCarthy, John J., and Cohn, Abigail (1998). Alignment and parallelism in Indonesian phonology. *Working Papers of the Cornell Phonetics Laboratory* 12. 6. Retrieved November 1, 2020 from https://scholarworks.umass.edu/linguist_faculty_pubs/6.
- Prince, Alan, and Paul Smolensky (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University, New Brunswick, New Jersey and University of Colorado, Boulder, Colorado.