# 國立政治大學語言學研究所碩士學位論文 

National Chengchi University
Graduate Institute of Linguistics
Master Thesis

## 從優選理論分析臺灣華語的韻母音組限制

An Optimality Theory Approach to Phonotactic Restrictions of
Taiwanese Mandarin Rhymes

指導教授：蕭宇超
Advisor：Yuchau E．Hsiao

研究生：葉德偉 撰
Student：Tei－wei Yeh

中華民國 —○九 年 六 月
June， 2020

# An Optimality Theory Approach to Phonotactic Restrictions of Taiwanese Mandarin Rhymes 



June, 2020


Copyright © 2020
Tei-wei Yeh
All Rights Reserved

## ACKNOWLEGEMENTS

感覺像是完成了一件人生大事，雖然過程很辛苦，但是看到成品時的成就感讓我覺得一切都是值得的。能多順利地完成這本論文，首先一定要再三地感謝蕭宇超老師，對蕭老師的感謝已經不僅僅是單純對指導教授的感謝，而是對於老師在我寫論文的期間是如何地為了讓我能夠準時畢業而犧牲了許多自己的時間一次又一次地解決我的疑難雜症。謝謝林蕙珊老師和吴貞慧老師所有寶貴的建議讓我的論文更完整，也謝謝兩位老師幾次的舟車勞頓來擔任我的口試委員。

謝謝惠鈴學姐在行政上的支持與照顧，在我每次疏忽了重要程序的時候不厭其煩的提點我；謝謝怡臻在我論文初期碰到困難時給了我很重要的意見；謝謝子權從以前對我的學期報告給出的建議，如今都成了論文的一部分；謝謝旺楨在工作室的行政支持讓我可以安心寫論文；謝謝欣蓉在我兩次口試的時候都主動幫助我處理一些雜務；謝謝仲良在最關鍵的時刻提供我重要的語料。

這段時間也非常感謝所有身邊朋友的陪伴。謝謝俊賢長期以來給我的歸屬感跟温暖；謝謝孟群，思妘和柏宇在我需要說話的夜裡用你們的聲音陪伴我；謝謝政矩和姿羽擔任我的心靈導師開導我；謝謝珮辰常在我需要放鬆的時候找我一起取暖；謝谢泰叡總是在我需要人陪的時候和我一起消磨時光；也謝謝徐臨嘉老師在我壓力最大的時候給了我精神上的支持，讓我可以渡過難關。

最後謝謝我的家人，不管我做什麼樣的決定，都無條件地支持我，讓我可以沒有後顧之憂地完成碩士學位。最重要的，我要謝謝我自己，雖然在整個碩士生涯中曾經幾度想直接放棄算了，但要不是我後來還是撐過來了，今天就不會有這様的感動。這段時間有太多不同的事情同時發生在我的生命中，說實在的，走得很辛苦，可是當我咬緊牙關走到今天，再回首一看，才發現，原來這就是人生精彩的所在。

## VITA

## Education

M.A. in Graduate Institute of Linguistics,

National Chengchi University
2017
B.A. in Department of English,

National Chengchi University

## Grants and Scholarships

Graduate Student Scholarship
National Chengchi University
Ministry of Science and Technology Scholarship
Project No. MOST 107-2410-H-004-115-MY2
2018-2020

2017
Graduate Student Scholarship
National Chengchi University

## Publication

"On Some Irregular Patterns of Syllable Contraction in Taiwan Mandarin." Paper presented at the $52^{\text {nd }}$ International Conference on Sino-Tibetan Languages and Linguistics, The University of Sydney, Sydney, Australia, June 26

## TABLE OF CONTENTS

## ACKNOWLEDGEMENTS $v$

VITA vi
TABLE OF CONTENTS vii
ABSTRACT ix
Chapter 1 Introduction 1
1.1 Preliminaries 1
1.2 Field of Research 2
1.2.1 The Domain of Rhyme 2
1.2.2 Dialect 3
1.3 Thesis Layout 5

Chapter 2 Theoretical Background and Literature Review 7
2.1 Optimality Theory 7
2.1.1 Constraint Ranking
2.1.2 Categories and Definitions of the Constraints 10
2.1.3 Local Conjunction and Self Conjunction 11
2.2 Previous OT Analyses of Mandaring C 12
2.2.1 The Issue of Mid Vowel Assimilation 12
2.2.2 The Issue of Low Vowel Raising 17

Chapter 3 Patterns of Rhymes in Taiwanese Mandarin 19
3.1 Principles and Charts 19
3.2 Controversial Rhyme Structures 21
3.2.1 [e] and [o] 21
3.2.2 The Low Front Vowel [a] 22
3.2.3 [jaj] 22

## Chapter 4 An OT Analysis 28

4.1 Rhymes without a Nasal Coda 28
4.1.1 Mid/Low Vowel Assimilation 28
4.1.2 The Phonotactic Restriction of Triphthongs 34
4.1.3 The Sequence of High Vocoids 40
4.1.4 Summary 46
4.2 Rhymes with a Nasal Coda 47
4.2.1 High Vowel Preceding a Nasal Coda 47
4.2.1.1 The Regular Patterns [in], [wən] and [yn] 48
4.2.1.2 The Irregular Outputs of $/ \mathrm{in} /$, /un/ and $/ \mathrm{yn} / \mathrm{L}$
4.2.2 [ən] and [əŋ] without a Prenuclear Glide 62
4.2.3 Low Vowel Preceding a Nasal Coda 64

4.2.4 Summary 74

Chapter 5 Conclusion 77
REFERENCES 80

國 立 政 治 大 學 研 究 所 碩 士 論 文 提 要

研究所別：語言學研究所
論文名稱：從優選理論分析臺灣華語的韻母音組限制
指導教授：蕭宇超
研究生：葉德偉
論文提要内容：（共一冊，16，505 字，分五章）

華語的音節結構非常簡單，最多僅能容納 CGVX 四個音段（Duanmu，2007）。然而其内部卻有著非常複雜的音組限制。先前的研究（Duanmu，2007；Lin， 2007／2015；Hsieh，2012）曾透過優選理論分析中元音的同化及低元音的提高現象，然這些研究並未針對臺灣口音中不同的特色做出考量；再者，除了這些音韻規則以外，華語的韻母仍存在許多其他的限制

本論文以臺灣的口音為基礎，透過優選理論（Prince and Smolensky，1993）分析華語的韻母音組限制，包含無鼻音韻尾及有鼻音韻尾雨個部分。無鼻音韻尾的韻母主要的現象包括中／低元音在雙元音結構下的同化，介音與滑音韻尾的異化及介音後不得緊接高元音；含鼻音韻尾的韻母結構中，主要的現象包含高／低元音與鼻音韻尾的同化，央元音的插入及高元音的提高。除了這些現象外，臺灣口音包含了少數不規則的韻母，如［iv］及［oy］，本論文認為此方言擁有較高排序的制約排除［əๆ］的組合（＊əๆ）及連續的後高音段（OcP－［＋hi，＋bk］）。

經過本論文的分析，臺灣華語的韻母中，大部分的空缺皆具系統性，僅＊［पa］及＊［ Cap$]$ ］無其他對稱的空缺，故本論文將此二韻母視為偶然空缺。而除了韻母内部外，聲母及韻母的組合仍存在著其他音組限制，則有賴後續更深入的研究。

## ABSTRACT

The syllable structure of Mandarin is quite simple which is described as CGVX (Duanmu, 2007), indicating that it can only contain four segments at most. However, there are complex restrictions within such simple structure. Previous research (Duanmu, 2007; Lin, 2007/2015; Hsieh, 2012) has proposed some analyses regarding to mid vowel assimilation and low vowel raising in terms of Optimality Theory. However, these studies do not focus on Taiwan dialect, which actually has some different rhymes from other dialects. Furthermore, there are still other phonotactic restrictions in addition to mid vowel assimilation and low vowel raising.

This thesis proposes an Optimality Theory (Prince and Smolensky, 1993) approach to analyze the phonotactic restrictions of Taiwanese Mandarin rhymes. The analysis is divided into two parts: One discusses the rhymes without a nasal coda and the other the rhymes containing a nasal coda. The phenomena in the rhymes without a nasal coda mainly include the mid/low vowel assimilation, the dissimilation of prenuclear and postnuclear glides and the prohibition of two adjacent high vocoids. As for the rhymes with a nasal coda, the analyses mainly include high/low vowels assimilating to the nasal coda, the central vowel insertion and low vowel raising. In addition to these phenomena, there are some irregular rhymes in Taiwanese Mandarin such as [iy] and [on]. These rhymes are surfaced out because of the high-ranked constraints such as *әy and $\mathrm{Ocp}_{\mathrm{cp}}[+\mathrm{hi},+\mathrm{bk}]$.

Through the analysis, most absent rhymes in Taiwanese Mandarin is systematic except for *[ча] and *[чач]. Therefore, they are considered accidental gaps in this thesis. In addition to the phonotactic restrictions within the rhyme domain, there are much more constraints regarding to the combination of the onset and the rhyme, which relies on further studies.

## Chapter 1

## Introduction

### 1.1 Preliminaries

The syllable structure of Mandarin is quite simple, which is composed of only four segments at most. However, many syllables that seem to be able to exist in Mandarin fail to surface out. That is, there are in fact complex phonotactic restrictions within Mandarin syllables.

Duanmu (2007) proposes the structure of CGVX to describe a Mandarin syllable, where $\underline{\mathbf{C}}$ represents a consonant, $\underline{\mathbf{G}}$ a glide, $\underline{\mathbf{V}}$ a vowel and $\underline{\mathbf{X}}$ a nasal coda or an offglide. This model indicates that not only is a Mandarin syllable composed of only at most four segments, the segment of each position also has a narrow inventory.

Traditionally, a Mandarin syllable is divided into two parts: Onset (the domain of C) and Rhyme (the domain of $\underline{\mathbf{G V X}})^{1}$. A lot of restrictions can be observed solely within the rhymes. For example, *[jan], *[wow] and *[e] $]^{2}$ are all absent in Mandarin even though $[\mathrm{j}, \mathrm{w}]$ are allowed to appear in the position of $\mathbf{G},[\mathrm{a}, \mathrm{o}, \mathrm{e}]$ are allowed to appear in the position of $\underline{\mathbf{V}}$ and $[\mathrm{n}, \mathrm{w}]$ are allowed to appear in the position of $\underline{\mathbf{X}}^{3}$.

[^0]Although some of the inexistent rhymes can be considered accidental gaps, most of the absent patterns seem to be systematic. For example, no diphthongs ${ }^{4}$ can follow the glide $[\Psi]$ to form a triphthong. Such systematic gaps indicate that there must be some rules and restrictions confining the combinations of Mandarin Rhymes.

Hence, this thesis aims to figure out all the systematic restrictions within Mandarin rhymes, based on Taiwan accent, in terms of Optimality Theory, and establish a complete constraint ranking system.

### 1.2 Field of Research

### 1.2.1 The Domain of Rhyme

The domain of rhyme in this thesis includes the prenuclear glide, the vowel, and the nasal coda or the off-glide. Specifically, a glide-initial syllable is viewed as a zeroonset syllable in this thesis, where the whole syllable is its rhyme, for it is observed that all glide-initial syllables in Mandarin can be preceded by another consonant to form a new syllable, as the examples shown in (1).
(1) Examples of glide-initial and true-consonant-initial syllables
a. yan $[\mathrm{jen}]$ 'eye' $\quad \operatorname{tian}\left[\mathrm{t}^{\mathrm{h}} \mathrm{jen}\right]$ 'to lick'
b. yao [jaw] 'medicine' miao [mjaw] 'temple'
c. wen $[\mathrm{w} \partial \mathrm{n}]$ 'warm' tun $\left[\mathrm{t}^{\mathrm{h}} \mathrm{w}\right.$ 'n] 'to swallow'

[^1]d. wei [wej] 'tiny' sui [swej] 'to follow'
e. yong [joy] 'servant' xiong [cjoy] 'elder brother'
f. yие [че] 'approximate' xие [єче] 'boot'

That is, a true-consonant-initial syllable and a glide-initial syllable only differ in whether there is a consonant filling in the $\underline{\mathbf{C}}$ slot. To sum up, only syllable-initial consonants are not included in the discussion. The domain that is discussed in this thesis is illustrated in (2).
(2) Illustration of the domain of rhyme $(\mathrm{O}=$ Onset, $\mathrm{R}=$ Rhyme $)$
a.

$t^{\text {h }}$ j $\quad \mathrm{e}$
qian 'thousand'

b. | O |  |  | R |
| :--- | :--- | :--- | :--- |
| ${ } }$ | G | V | X |

j e n
yan 'cigarette'
c.

d.

a $n$
an 'safe'

### 1.2.2 Dialect

Mandarin is widely spoken all over the world, and therefore there are a lot of different dialects and accents. To specify, we will analyze Mandarin rhymes based on Taiwan accent. The different rhymes between Taiwan accent and Beijing accent are listed in (3).
(3) Special rhymes in Taiwanese Mandarin

| TM rhyme | Example | Gloss | Beijing accent |
| :---: | :---: | :---: | :---: |
| a. $\partial$ | [kə] | ge 'song' | [kr] |
| b. in | [6iy] | xing 'star' | [¢jə刀] |
| c. oy | [(w)oy] <br> [toy] | weng 'old man' tung 'winter' | $\begin{aligned} & {[\mathbf{w} \boldsymbol{\eta}]} \\ & {[\mathrm{toy}]} \end{aligned}$ |

In Taiwan accent, a mid vowel appearing in the rhyme domain alone surfaces as [ə], which is similar to the mid central vowel preceding a nasal coda, illustrated in (3a). The spectrograms in $(4-5)$ produced by the praat serve to support this observation.
(4) The spectrograms of Taiwan accent
a. Lengqqi 'air conditioner' $(\mathrm{F} 2=1396) \quad$ b. Tese $\boldsymbol{e}^{\prime}$ characteristic' $(\mathrm{F} 2=1399)$

(5) The spectrograms of Beijing accent


In (4), the F2's, which indicate the backness of a vowel, are close, which means that the vowels in the syllables leng and se are the same in Taiwan accent. On the other hand, in (5), the F2's of the vowels in leng and se have a great distance, meaning that they are two different vowels.

As for (3b), some may argue that [in] and [in] have been neutralized in Taiwan accent. In spite of this fact, both [in] and [in] are still heard in the dialect even though speakers may not choose the "correct" coda. For example, both xin 'new' and xing 'star' are able to be pronounced as [cin] or [ciy]. Such phenomenon of neutralization is also seen when it comes to the rhymes [ən] and [əๆ] (Fon et al, 2011).

In Beijing dialect, the rhyme [on] in (3c) must follow a consonantal onset. Otherwise, it will be pronounced as [wəy] (Lin, 2007), as shown in (3). However, there is no such distinction for the rhyme [on] in Taiwan dialect. Though sometimes Taiwanese Mandarin speakers pronounce the world weng as [won], it is undoubted that in this case the initial [w] is syllabified to the onset position since [woy] cannot follow any consonant. Therefore, the rhyme part of this syllable is still [on].

In addition to the differences listed above, there is still another special part worth mentioning. That is, the $-r$ suffix is absent in Taiwanese Mandarin. Therefore, the $-r$ suffix is not going to be discussed. On the other hand, the rhyme [ $\mathfrak{\gamma}$ ] (pinyin: er, Bopomofo: $ル$ ) do exist in Taiwan dialect. However, it can only solely appear in a syllable without any other segments. Due to its special phonological status, this rhyme will not be discussed, either.

### 1.3 Thesis Layout

This thesis consists of five chapters. Chapter One introduces the basic background of Mandarin syllable structure and Mandarin rhymes. Afterwards, we also specify what
is going to be discussed in this thesis. In Chapter Two we introduce some theoretical backgrounds and review previous research on Mandarin phonotactic restrictions. Chapter Three lists the patterns of Taiwanese Mandarin rhymes and generally describes the patterns. Main analysis lies in Chapter Four. We first analyze the rhymes without nasal codas, and then go on to the rhymes with a nasal coda. The last chapter sums up the constraint ranking arranged in Chapter Four and ends up with further issues.


## Chapter 2

## Theoretical Background and Literature Review

### 2.1 Optimality Theory

Different from traditional rule-based phonology, Optimality Theory (henceforth OT), first developed by Prince and Smolensky (1993), emphasizes that it is a string of constraints that confines the surface forms of a certain language. A typical example proposed by Kager (2004) is illustrated in (1) below.
(1) /sænd/ $\rightarrow$ [sæ̃nd] (Kager, 2004)

| /sænd/ |  | $* V_{\text {oral }}$ | $* V_{\text {NASAL }}$ |  |
| :--- | :---: | :---: | :---: | :---: |
| a. sænd |  | $*!$ |  | Ident-IO(nasal) |
| b. $\rightarrow$ sãnd |  |  | $*$ |  |

The constraints can be paraphrased as in $(2-4)$, in terms of McCarthy's (2008) definition formula.
(2) *Voral $N$

Assign one violation mark for every oral vowel preceding a nasal segment.
(3) $* \mathrm{~V}_{\mathrm{NASAL}}$

Assign one violation mark for every nasalized vowel.
(4) Ident-IO(nasal)

Assign one violation mark for every segment in the output that has a different value for the [nasal] feature from its correspondent input.

### 2.1.1 Constraint Ranking

As is shown in (1), there is a certain order for the string of constraints, which is called constraint ranking. Constraints ranked higher are less violable than those ranked lower. In example (1), the highest ranked constraint is $* V_{\text {oraL }} N$, and then * $\mathrm{V}_{\text {nasal }}$. The lowest one is Ident-IO(nasal). Such constraint ranking is represented as below:

$$
* V_{\text {ORAL }} \gg * V_{\text {NASAL }} \gg \text { Ident-IO(nasal) }
$$

If the constraint is violated, an asterisk is marked in the tableau. An exclamation mark following the asterisks indicates that the candidate is ruled out by that constraint. Let's take a look at the example in tableau (1). Candidate (a) [sænd] violates the constraint $* V_{\text {ORAL }} N$ since there is an oral vowel $[æ]$ preceding the nasal $[n]$. Therefore, an asterisk is marked to indicate the violation. Likewise, candidate (b) [s̃̃nd] violates the other two constraints since there is a nasal vowel $[\tilde{\mathfrak{x}}]$ in the output and the vowel has different values for the [nasal] feature between its input and output. Asterisks are therefore marked in the correspondent boxes.

For these three constraints, Candidate (a) only violates one constraint while candidate (b) violates two. However, *VoraLN, which is violated by candidate (a) and obeyed by candidate (b), is ranked the highest among the three; therefore, candidate (a) is first ruled out. An exclamation mark is thus added after the asterisk. The arrow preceding candidate (b) indicates that it is the "winner." The shadowed boxes indicate that the constraints are not active because the winner has been selected.

The constraint ranking can also be illustrated as a Hasse Diagram (McCarthy, 2008) as (5) below.
(5) The Hasse Diagram of the example

In diagram (5), the vertical lines indicate domination, where we can interpret this diagram as "* $V_{\text {ORAL }} \mathrm{N}$ dominates $* V_{\text {NASAL }}$, and $* V_{\text {NASAL }}$ dominates Ident-IO(nasal)." There might be some cases where there are two or more constraints ranked the same high.

Different constraint domination relations are illustrated in (6) below. (6a) indicates that both Max and Dep dominate Onset, but there is no domination relation between Max and Dep. (6b) indicates that *V\# dominates both Ident(long) and Max, but there is no domination relation between $\operatorname{Ident}($ long ) and Max. (6c) indicates that both Ident(long) and Max are dominated by *Comp-Syll, * ${ }^{\text {unsyll }}$ and Max-C simultaneously, but there is no domination relation between $\operatorname{Ident}($ long $)$ and $\mathrm{Max}_{\text {a }}$ and among *CompSyll, * $C^{\text {unsyll }}$ and Max-C.
(6) Hasse Diagrams of different constraint domination relations (McCarthy, 2008)
a. Max, Dep >> Onset

b. *V\# >> Ident(long), Max
c. *Comp-Syll, *Cunsyll, Max-C $\gg$ Ident(long), Max

2.1.2 Categories and Definitions of the Constraints

To take a view of the definitions of constraints $(2-4)$, it is found that constraint (4) confines the output with a different aspect from constraints (2) and (3). Constraints (2) and (3) define what should a surface form of this language be like, which have nothing to do with the input. This category of constraints is called markedness constraint. Markedness constraints do not mention the input form or the input-output mapping. They focus on the presence of the output form (McCarthy, 2008).

Constraint (4) mentions the input-output mapping, which is different from constraints (2) and (3). This category of constraint is called faithfulness constraint. McCarthy and Prince (1995/1999) propose the correspondence theory, which states that each candidate that is evaluated supplies an output form that has correspondence
relations with the input. Faithfulness constraints focus on the correspondence relations between the input and the output.

In some situations, one constraint can be violated not only once by a single candidate. An example proposed by McCarthy (2008) is shown in (7).
(7) $/ \mathrm{bad} / \rightarrow[$ bat $]$ McCarthy, 2008)

| $/$ bad $/$ | *Voice | Ident([voice]) |
| :--- | :---: | :---: |
| a. $\rightarrow$ bat | $*$ | $*$ |
| b. bad | $* *!$ |  |

In tableau (7), constraint *VoIce, prohibiting voiced consonants to surface in the output, is violated by candidate (b) twice because [b] and [d] are both voiced. However, it is necessary to carefully define how violation marks should be given. Think that if we consider that constraint *Voice directly rules out candidates that have voiced consonants, the winner of example (7) will be candidate (b) because both candidates violate *Voice but only candidate (a) violates Ident([voice]). The correct output is not predicted in this case. To avoid this situation, McCarthy (2008) advises that a constraint be defined starting with "Assign one violation mark for every..." This thesis also adopts this suggestion to define every active constraint in the paper.

### 2.1.3 Local Conjunction and Self Conjunction

In OT, two constraints can be conjoined together to form a local conjunction. When a local conjunction conjoins constraint A and constraint B , it is violated only if the candidate violates both constraints A and B in a given domain $\delta$. For simplification,
a local conjunction conjoining constraint A and constraint B in domain $\delta$ is noted as $\{A \& B\}_{\delta}$ in this thesis. Universally, $\{A \& B\}_{\delta}$ is ranked higher than both $A$ and $B$.

Self-conjunction, like local conjunction, is a conjoined constraint which is violated when a certain constraint is violated twice in a given domain. Tableau (8) is an example of self-conjunction proposed by Ito and Mester (1998). A self-conjunction of constraint A in domain $\delta$ is noted as $\mathrm{A}^{2}{ }_{\delta}$. Like local conjunction, $\mathrm{A}^{2}{ }_{\delta}$ is universally ranked higher than A, usually with another constraint ranked in between.
(8) $/ b^{\mathrm{h}} \mathrm{id}^{\mathrm{h}} / \rightarrow\left[\mathrm{b}^{\mathrm{hid}}\right]$ or $\left[\mathrm{bid}^{\mathrm{h}}\right]$ (hypothetical) (Ito and Mester, 1998) ${ }^{5}$


### 2.2 Previous OT Analyses of Mandarin

### 2.2.1 The Issue of Mid Vowel Assimilation

Previous research has done some analyses about the phonotactic restrictions of Mandarin. Duanmu (2007) has analyzed the phonotactic restrictions of Mandarin rhymes through OT. In his analysis, he derives two constraints to evaluate Mandarin rhymes: NC-Harmony and GN-Harmony. The former indicates that the nucleus and the

[^2]coda of a same syllable should share the same value for [back] and [round] feature. The latter one indicates that the nucleus and the prenuclear glide of a same syllable should share the same value for [back] and [round] feature. Between the two constraints, NCHarmony is ranked higher than GN-Harmony. The examples by Duanmu are shown in (9) and (10) below.
(9) /чә/ $\rightarrow$ [че] (Duanmu, 2007)

| /чә/ | NC-Harmony | Avoid-[ø] | \{GN-[back], GN-[round]\} |
| :--- | :---: | :---: | :---: |
| а. чә |  |  | $*$ |
| b. $\rightarrow$ че |  |  | $*$ |
| c. чø |  | $*!$ |  |

(10) /wəj/ $\rightarrow$ [wej] (Duanmu, 2007)

| $/$ wəj $/$ | NC-Harmony | GN-Harmony |
| :--- | :---: | :---: |
| a. wəj | $*!$ | $*$ |
| b. $\rightarrow$ wej |  | $*$ |
| c. woj | $*!C h$ |  |

In tableau (9), the input /чə/ has no coda, so NC-Harmony is not violated by any candidates. This situation is called vacuous satisfaction, and the definition is in (11) below.
(11) The definition of "vacuous satisfaction" (McCarthy, 2008)

When a candidate does not contain a structure that a certain constraint is against, it is stated that the constraint is vacuously satisfied by the candidate. For instance, an open syllable does vacuously satisfy a constraint requiring the coda to be voiceless. In OT analysis, vacuous satisfaction is treated the same as nonvacuous satisfaction, though.

Let's get back to example (9). All candidates satisfy NC-Harmony, so none is ruled out. Among the three candidates, candidate (c) completely satisfies GN-Harmony, but it contains a front rounded vowel [ø], which is not licensed in Mandarin. Therefore, it is ruled out by a higher-ranked constraint Avoid-[ø]. As for candidates (a) and (b), (a) violates GN-Harmony twice because the prenuclear glide and the nucleus has different values for both [back] and [round] features while (b) only violates GN-Harmony once, with its prenuclear glide and nucleus only differing in [round] feature. Therefore, candidate (b) is the best choice for the output form. Note that Avoid- $[\varnothing]$ must be ranked higher than GN-Harmony, or candidate (c) will be wrongly predicted as the winner.

Example (10) illustrates how NC-Harmony is ranked higher than GN-Harmony. Comparing candidates (b) and (c), the former violates GN-Harmony while the latter NC-Harmony. The attested output is (b) [wej], so NC-Harmony is less violable.

Lin (2015) also analyzes the mid vowel assimilation phenomenon of different Mandarin dialects with OT approach. The research states that different dialects have different constraint rankings. Among them, the constraint ranking of Standard Mandarin is shown in (12).
(12) The constraint ranking about mid vowel assimilation of Standard Mandarin (Lin, 2015) ${ }^{6}$
*ø >> Rime-Harmony, GV-Harmony >> *o >> *e

Likewise, Lin ranked the constraint * $\varnothing$ on the top to avoid such segment. Although Lin accidentally ranked Rime-Harmony and GV-Harmony the same, the used constraints are almost identical to those used by Duanmu. Both of them high-rank a constraint to restrict that $\underline{\mathbf{V}}$ and $\underline{\mathbf{X}}$ in a same syllable have identical values for the [back] (and also [round]) feature. However, it is observed that there are different phenomena when the $\underline{\mathbf{X}}$ slot is filled with a nasal coda or an off-glide, as is illustrated in (13).
(13) The mid vowel assimilation when $\underline{\boldsymbol{X}}$ is a nasal coda or an off-glide

| $/ \mathrm{j} / \rightarrow[\mathrm{ej}]$ | $/ \mathrm{a} /$ becomes [e] because of the [-back] feature of [j], which satisfies NC-Harmony. |
| :---: | :---: |
| /ən/ $\rightarrow$ [ən] | /2/remains as a central vowel despite the following [-back] nasal [n]. |
| /əw/ $\rightarrow$ [ow] | $/ 2 /$ becomes [0] because of the [+back] feature of [w], which satisfies NC-Harmony. |
| /əŋ/ $\rightarrow$ [əŋ] | /2/ remains as a central vowel, which still satisfies NC-Harmony because of the following [+back] nasal [ y$]$. |

According to (13), it is assumed that the mid vowel has a different requirement of assimilation to a nasal coda from that to an off-glide. Therefore, it is necessary to re-

[^3]rank the constraints or introduce new ones in order to fulfill the two different phenomena in table (13).

Lin (2007) has also discussed about the difference of the mid vowel assimilation when it is preceded by a nasal coda or an off-glide. The problem does not lie in the segment filled in the $\underline{\mathbf{X}}$ slot itself because it is observed that low vowels also assimilate to nasal codas. For example, shan [san] 'mountain' and shang [say] 'injury' are allowed, but *[say] and *[san] are not. Therefore, it is unreasonable to argue that nasal codas do not trigger the assimilation. To solve this problem, Lin proposes a hypothesis: there are two main kinds of VX structure in Mandarin; one is diphthong and the other is "high/low vowel + nasal." She excludes the structure of "mid vowel + nasal" in the underlying form. In her theory, a mid vowel only precedes a nasal coda due to the condition and rule described in (14). There are no mid vowels preceding a nasal coda in the underlying forms.
(14) The [2]-insertion condition
a. Environment

When a high vowel and its following nasal coda have different values for the [back] feature, an [ $\partial$ ] is inserted between them.
b. Rule

$$
\emptyset \rightarrow \partial /\left[\begin{array}{c}
-\mathrm{cons} \\
\alpha \mathrm{bk}
\end{array}\right]-\left[\begin{array}{c}
+ \text { nas } \\
\beta \mathrm{bk}
\end{array}\right]
$$

c. Examples (The glide formation rule is omitted.)
(i) /lun/ $\rightarrow[\mathrm{lw} 2 \mathrm{n}]^{\text {' }}$ wheel'
(ii) $\quad / \mathrm{piy} / \rightarrow[\mathrm{pj} \partial \mathrm{y}]$ 'ice'

Lin's hypothesis seems to successfully explain why mid vowels do not assimilate to nasal codas. However, she does not clearly explain why there are syllables like shen [sən] 'body' and seng [səy] 'monk', where there are no high vocoids in the syllables to trigger the [ə]-insertion rule.

Another problem lies in the example (ii) in (14c). In Taiwan accent, the word bing 'ice' is pronounced as [pin]. Other words that have the same rhyme are also pronounced this way. This pronunciation violates Lin's Rime-Harmony constraint in (12) and her [ə]-insertion rule in (14) because [i] is [-back] while [y] is [+back]. Therefore, it is necessary to re-rank the constraints or propose new ones to fulfill the dialect spoken in Taiwan.

### 2.2.2 The Issue of Low Vowel Raising

Low vowel raising in Mandarin is a special phonological phenomenon that it is hard to find another similar rule in other contexts. The simplified formula is illustrated in (15).
(15) The low vowel raising rule

$$
\mathrm{a} \rightarrow \mathrm{e} / \mathrm{j}, \mathrm{u} \ldots \mathrm{n}
$$

Aiming at this phonological phenomenon, Hsieh (2012) has analyzed it via rulebased phonology and OT. In his research, he proposes a self-conjunction constraint called "Agree $[\text { back }]^{2 "}$ " to rule out the ill-formed *[jan] and *[yan], and successfully predicts the well-formed [jen] and [yen]. The definition of the constraint is shown in (16).
(16) Agree $[\text { back }]^{2}$ (Flemming, 2003; Hsieh, 2012)

Assign one violation mark for every syllable only if where there is a vowel not having the same value for the [back] feature as both of its adjacent consonants.

In Hsieh's research, constraint (16) rules out *[jan] but is satisfied by [jen]. He follows the SPE tradition to regard the low front vowel [a] as [+back]. This theory does support constraint (16) to rule out *[jan] because [a] is [+back] while [j] and [n] are [back]. However, if we specify [a] as [+back], other problems might be caused, as elaborated below.

First, the mid vowel assimilation and low vowel assimilation in diphthongs are no longer the same issue. As what is mentioned in section 2.2.1, the nucleus must have the same value for the [back] feature as the off-glide. The fact is that we have the rhyme [aj] (cf. [ej]) in Taiwanese Mandarin, which obviously violates the mentioned constraint if [a] is regarded as [+back]. Second, it will be difficult to explain why a low vowel is surfaced "even more retracted" as [a] when preceding [+back] codas like [w] and [ y ] to form rhymes like [aw] and [aŋ] because [a] is already specified as [+back]. In other words, now that both [a] and [a] are both [+back], then what feature distinguishes these segments?

Due to these problems, it is not suitable to specify [a] as [+back] because it will lead to more obstacles for analysis, and the constraint Agree[back] ${ }^{2}$, therefore, cannot rule out the ill-formed *[jan] and *[yan]. In order to take other phonological phenomena (e.g. [aj] and [ej] are composed of a front vowel and [j]; the two segments agree with each other in [back] in rhymes like [aw] and [ay]) into consideration, a new constraint instead of AGREE[back $]^{2}$ should be proposed for low vowel raising.

## Chapter 3

## Patterns of Rhymes in Taiwanese Mandarin

### 3.1 Principles and Charts

This section lists all the rhymes that are used in Taiwanese Mandarin. The inventories of each position in the GVX structure are listed in (1).
(1) The list of phonetic segments for each position
$\underline{\mathbf{G}} \quad \emptyset \quad \mathrm{j} \quad \mathrm{w} / \mathrm{u}$
$\underline{V}(i, u, y) \quad a \quad a \quad e \quad a \quad o$
X $\quad \varnothing \quad \mathrm{j} \quad \mathrm{w}-\mathrm{n}$

As (1) shows, there are three glides can fill in the $\underline{\mathbf{G}}$ slot, namely $[j],[\mathrm{w}]$, and $[\mathrm{u}]$. Since a syllable can have no prenuclear glide, this position can also be filled with a $\emptyset$. There are multiple choices for the $\underline{\mathbf{V}}$ slot. Among the high vowels [i], [u] and [y], only the one correspondent to the prenuclear glide $([j]$ to $[\mathrm{i}],[\mathrm{w}]$ to $[\mathrm{u}]$ and $[\mathrm{u}]$ to $[\mathrm{y}])$ can fill in the $\underline{\mathbf{V}}$ slot. As for the low vowels, $[\mathrm{a}]$ and [a] are derived from a same phoneme $/ \mathrm{a} /$. Mid vowels in the $\underline{\mathbf{V}}$ position is also this case. The underlying form of the mid vowel is hypothesized as $/ \partial /$. The $\underline{\mathbf{X}}$ slot can be filled with an off-glide $[\mathrm{j}] /[\mathrm{w}]$, or a nasal coda $[\mathrm{n}] /[\mathrm{n}]$. Like the $\underline{\mathbf{G}}$ position, the $\underline{\mathbf{X}}$ slot can also be filled with a $\emptyset$ to lead to an open syllable. Charts $(2-5)$ illustrate all the existent and inexistent rhymes in Taiwanese Mandarin.
(2) $\underline{\mathbf{G}}=\emptyset$

| $\underline{\mathbf{V}} \backslash \underline{\mathbf{X}}$ | $\emptyset$ | j | w | n | 1 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| e | *e | ej | *ew | *en | *en |
| $ə$ | $ə$ | * ${ }^{\text {j }}$ | *วw | әn | ə刀 |
| o | * | * ${ }_{\text {j }}$ | ow | *on | on |
| a | a | aj | *aw | an | *ay |
| a | *a | *aj | aw | *an | a) |

(3) $\underline{\mathbf{G}}=\mathrm{j}$

(4) $\underline{\mathbf{G}}=\mathrm{w}$

| $\underline{\mathrm{V}} \backslash \underline{\mathrm{X}}$ | $\emptyset$ | J | W | n | 1 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| u | u | *uj ${ }^{*}$ wi | $\mathrm{uw}(\mathrm{wu})=\mathrm{u}$ | * un | *uŋ |
| e | *we | wej | * wew | *Wen | *wey |
| $\bigcirc$ | *Wə | * wəj | *Wəw | wən | *Wəŋ |
| 0 | wo | * woj | * wow | *Won | *Woy |
| a | *wa | waj | *waw | wan | *way |
| a | wa | * waj | * waw | *wan | way |

(5) $\underline{\mathbf{G}}=\mathrm{u}$

| $\underline{\mathbf{V}} \backslash \underline{\mathbf{X}}$ | $\emptyset$ | j | w | n | ๆ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| y | y |  | $*_{\mathrm{yw}}=* \mathrm{yu}$ | yn | *y |
| e | че | * чеј | * 4 ew | чen | * чеп |
| $\partial$ | * чә | * чəј | * पәW | * цəп | * чəŋ |
| 0 | * чо | * чој | * Yow | * yon | * чо才 |
| a | * ¢а | * ¢ај | * yaw | * ¢ ${ }^{\text {an }}$ | * ча, |
| a | * чa | * чaj | * yaw | * yan | * yaŋ |

Charts $(2-5)$ list all the present and absent rhymes in Taiwanese Mandarin. However, some inexistent rhymes might be controversial, which will be clarified in the next section.

### 3.2 Controversial Rhyme Structures

### 3.2.1 [e] and [o]

The mid vowels [e] and [ o ] (see the leftmost column in chart (2)) only appear in interjections or particles in Taiwanese Mandarin no matter they are in onsetless syllables or follow a consonant. Never do these rhymes appear in lexical words. Therefore, the surface of the rhymes [e] and [o] is viewed as exceptions in this thesis, which are still considered inexistent.

### 3.2.2 The Low Front Vowel [a]

According to Lin (2007), the low vowel [a] is considered [-back] in Mandarin since there are only two low vowels, [a] and [a], in the patterns, and [a] is more advanced than [a]. When there is no other segment in the rhyme domain except for a low vowel, the front [a] is surfaced (e.g. $b a$ [pa] 'eight'; sha [sa] 'sand'), as is shown in the leftmost column in chart (2). Therefore, the low front vowel [a] is viewed as the underlying form of the low vowel.

### 3.2.3 [jaj]

The rhyme [jaj] in chart (3) is considered absent in Taiwanese Mandarin. Some may argue that this rhyme appears in the word yai [jaj] 'cliff'. However, this is the only word that contains this rhyme. Hsieh (2012) states that Mandarin speakers tend to pronounce this word as [ja] in casual speech while Duanmu (2007) thinks people pronounce this word as either [ja] or [aj].

In order to confirm how Taiwanese Mandarin speakers pronounce this word in casual speech, a simple experiment has been done. Six native speakers were asked to read several sentences casually, and one of them contained the word yai. Consequently, five of the participants pronounced the word as [aj], and the other pronounced it as [jaj], but further admitted that this is actually a word hard to pronounce when the conductor made sure whether she really pronounced this way in casual speech.

Though this experiment, different from what has been proposed by Hsieh and Duanmu, does not attest the output form of [ja], it proves that [jaj] can be reasonably viewed as an exception.

### 3.3 A General Description

According to previous analyses (Hashimoto, 1970; Duanmu, 2007; Lin, 2007), there are only one mid vowel phoneme and one low vowel phoneme in Mandarin. That is, different mid vowels in Mandarin are derived from a same phoneme in the underlying form, and so are different low vowels. With this theoretical knowledge, it can be assumed that a mid (low) vowel will not have minimal pairs with one another. For example, among [wej], *[waj] and *[woj] in chart (4), only one of them is existent because the three mid vowels should not have minimal pairs. Chart (6) labels the "phonemic group" with bold lines based on this theory. In (6), a block squared by bold lines should theoretically have only one legal pattern.
(6) Phonemic group indication chart ( $\underline{\mathbf{G}}=$ any $)$

| $\underline{\mathbf{V}} \backslash \underline{\mathbf{X}}$ | $\varnothing$ | j | w | n | p |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $(\mathrm{i}, \mathrm{u}, \mathrm{y})$ |  |  |  |  |  |
| e |  |  |  | 0 |  |
| $\partial$ |  | $C h e n$ | hi |  |  |
| o |  |  |  |  |  |
| a |  |  |  |  |  |
| a |  |  |  |  |  |

For example, if we apply chart (6) to the rhyme patterns without a prenuclear glide $(\underline{\mathbf{G}}=\emptyset)$, chart (7) is obtained as below. Theoretically, each bold-lined square should have only one pattern allowed. Therefore, among [ej], ${ }^{*}[\mathrm{jj}]$ and ${ }^{*}[\mathrm{oj}]$, only [ej] is allowed. Likewise, [aw] is allowed while *[aw] is not. If this rule is violated, then there
should be a problem to be dealt with. For example, in chart (7), both [əท] and [on] are allowed even though they are in the same bold-lined square. This irregular phenomenon thus deserves discussions.
(7) Phonemic group indication chart $(\underline{\mathbf{G}}=\varnothing)$

| $\underline{\mathbf{V}} \backslash \underline{\mathbf{X}}$ | $\emptyset$ | j | w | n | y |
| :---: | :---: | :---: | :---: | :---: | :---: |
| e | *e | ej | *ew | *en | *en |
| ә | $ə$ | * ${ }^{\text {j }}$ | * ${ }^{\text {W }}$ | әn | əท |
| o | * | * ${ }^{\text {j }}$ | ow | *on | or |
| a |  | aj | *aw | an | *ay |
| a | *a | *aj | aw | *an | ay |

However, this way of division is still too general. Remember that in section 2.2.1, it has been mentioned that $\operatorname{Lin}(2007)$ states that mid vowels preceding a nasal coda are inserted, whose underlying form is in fact a null, as is shown in (8). Therefore, the border between $\underline{\mathbf{V}}=\varnothing$ and the mid vowels in the nasal-coda columns should be broken. Chart (9) is the revision of (6), where the dotted lines indicate the broken ones.
(8) The [ə]-insertion rule

$$
\phi \rightarrow \partial /\left[\begin{array}{c}
-\mathrm{cons} \\
\alpha \mathrm{bk}
\end{array}\right]-\left[\begin{array}{c}
+\mathrm{nas} \\
\beta b k
\end{array}\right]
$$

(9) Phonemic group indication chart $(\underline{\mathbf{G}}=$ any $)$ (revised)

| $\underline{\mathbf{V}} \backslash \underline{\mathbf{X}}$ | $\emptyset$ | j | w | n | y |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $(\mathrm{i}, \mathrm{u}, \mathrm{y})$ |  |  |  |  |  |
| e |  |  |  |  |  |
| $\partial$ |  |  |  |  |  |
| o |  |  |  |  |  |
| a |  |  |  |  |  |
| a |  |  |  |  |  |

If we adopt the model in chart (9) to each chart in $(2-5)$, several counterexamples may appear:
(i) In chart (2), both [ər] and [oy] are allowed. The two mid vowels have minimal pairs.
(ii) In chart (3), no mid vowels can be put between two [j]'s. *[jej], *[joj] and *[joj] are all not allowed. This seems to indicate that a mid vowel cannot surface in such environment. The same case is found when it comes to the low vowels. Neither [a] nor [a] can appear between two [j]'s.
(iii) In chart (3), both [in] and [jen] are allowed. This seems to violate Lin's [ə]insertion rule (see 2.2.1). However, no low vowels can be between [j] and [n]. The similar case is found when it comes to the pair [iŋ] and [jon], whereas we have [jay] this time.
(iv) Like (ii), neither mid vowels nor low vowels can appear between to [w]'s in chart (4).
(v) In chart (4), no matter whether a mid vowel is inserted or not, *[un], *[wen], *[wəy] or *[woy] are all absent in the patterns. This phenomenon is also observed in
the patterns *[yŋ], *[чеп], *[чəŋ] and *[чоу] in chart (5).
(vi) Like (iii), both [yn] and [чen] are allowed. However, no low vowels appear between $[\mathrm{Y}]$ and $[\mathrm{n}]$ in chart (5). In fact, not only when there is a coda $[\mathrm{n}],[\mathrm{Y}]$ cannot precede any low vowels.
(vii) When there is a glide coda [j] or [w], neither mid vowels nor low vowels can surface after $[\Psi]$ in chart (5). That is, $[\Psi]$ cannot precede any diphthongs.
(viii) Two adjacent high vocoids like *[ju] in (3), *[wi] in (4), and *[ui] and *[uu] ${ }^{7}$ in (5) are all unallowed. To make it clearer, (10) lists these violations of the phonemic group division in a brief way.
(10) The list of violations of the phonemic group division
a. [əๆ] and [oŋ] are both allowed. (see (2))
b. $\quad *[j e j], *[j j j]$ and $*[j o j]$ are all not allowed. (see (3))
c. ${ }^{*}[\mathrm{jaj}]$ and $*[j a j]$ are both not allowed. (see (3))
d. [in] and [jen] are both allowed. (see (3))
e. *[jan] and *[jan] are both not allowed. (see (3))
f. [ii] and [joy] are both allowed. (see (3))
g. ${ }^{*}[w e w], *[w ə w]$ and $*[$ wow $]$ are all not allowed. (see (4))
h. *[waw] and *[waw] are both not allowed. (see (4))
i. *[up], *[wen], *[wəy] and *[woy] are all not allowed. (see (4))
j. $\quad *[ч а]$ and $*[ч a]$ are both not allowed. (see (5))

[^4]k. *[чеј], *[чәj] and *[чој] are all not allowed. (see (5))

1. *[чај] and *[чaj] are both not allowed. (see (5))
m. *[чеw], *[чəw] and *[чow] are all not allowed. (see (5))
n. *[yaw] and *[yaw] are both not allowed. (see (5))
o. [yn] and [yen] are both allowed. (see (5))
p. *[ $¢ \mathrm{an}]$ and *[yan] are both not allowed. (see (5))
q. *[yŋ], *[чеŋ], *[чəŋ] and *[чоŋ] are all not allowed. (see (5))
r. *[чаŋ] and *[чaŋ] are both not allowed. (see (5))
s. Each of *[ju], *[wi], *[qi] and *[yu] is not allowed. (see (2-5))


## Chapter 4

## An OT Analysis

In this chapter, an OT analysis will be displayed and divided into two sections. First we will discuss about rhymes without a nasal coda, including the following structures: $\underline{\mathbf{V}}, \underline{\mathbf{G V}}, \underline{\mathbf{V G}}, \underline{\mathbf{G V G}}$. The other section will focus on rhymes containing a nasal coda, including the structures of $\underline{\mathbf{V N}}$ and $\mathbf{G V N}$. For the structural indicating capital letters, the $\underline{\mathbf{G}}$ preceding $\underline{\mathbf{V}}$ indicates the prenuclear glides and the $\underline{\mathbf{G}}$ following $\underline{\mathbf{V}}$ indicates the postnuclear ones. For other symbols, $\underline{\mathbf{V}}$ stands for vowel and $\underline{\mathbf{N}}$ stands for nasal coda.

### 4.1 Rhymes without a Nasal Coda

### 4.1.1 Mid/Low Vowel Assimilation

As is mentioned in section 2.2.1, the phenomenon of mid/low vowel assimilation has been widely discussed in previous research (Duanmu, 2007; Lin, 2015). Let's review their constraints about this phonological phenomenon in $(1-2)$.
(1) NC-Harmony (Duanmu, 2007) or Rime-Harmony (Lin, 2015)

Assign one violation mark for every syllable whose nucleus and coda have different values for [back] and [round] features.
(2) GN-Harmony (Duanmu, 2007) or GV-Harmony (Lin, 2015)

Assign one violation mark for every syllable whose nucleus and prenuclear glide have different values for [back] and [round] features.

Both of these analyses mention the [back] and [round] features in a single constraint. However, in order to avoid predicting the wrong outputs *[чə], Duanmu assigns two violation marks for each of the two features, as (3) shows.
(3) $/$ чә/ $\rightarrow$ [че] (Duanmu, 2007)

| $/$ чә/ | NC-Harmony | Avoid-[ $\varnothing]$ | \{GN-[back], GN-[round]\} |  |
| :--- | :---: | :---: | :---: | :---: |
| a. чә |  |  | $*$ | $*!$ |
| b. $\rightarrow$ че |  |  |  | $*$ |
| c. чø |  |  |  | $*$ |

The analysis in (3) rules out *[чə] because [ч] and [ə] do not agree in not only [round] feature but also [back] feature. Therefore, I think it will be better if we separate the constraint into two, one assigning a violation mark for the [back] feature and the other for the [round] feature.

Moreover, in Duaumu's analysis, there is a high-ranked constraint "Avoid-[ø]" to prevent the ill-formed $*[\varphi \varnothing]$ from surfacing out. In Taiwanese Mandarin though, in addition to [ø], we still need to avoid segments like [ $\gamma$ ] or [p]. Therefore, we can just high-rank a constraint to avoid all the segments that are not listed in the inventory instead of introducing such specific constraints, which is displayed in (5). Constraints (6-9) are separated from Duaumu's and Lin's based on the different features. Constraint (10) is developed to ensure it is the vowel instead of the glide that changes
its feature. Constraint (11) is the violated faithfulness constraints. The descriptive generalization is in (4).
(4) The descriptive generalization for mid/low vowel assimilation

The vowel should agree with the prenuclear glide in [back] and [round] features, but if there is a glide coda following it, it should agree with the following glide. This requirement is enforced by vowel fronting/backing and rounding.
(5) Phonotactics

Assign one violation mark for every $\varnothing, ⿷, \amalg, \gamma, \mathrm{D}$ and $\underset{\mathrm{u}}{ }$ in the output.
(6) Agree[bk]-VG

Assign one violation mark for every syllable whose vowel and glide coda do not agree in the [back] feature.
(7) Agree[rd]-VG

Assign one violation mark for every syllable whose vowel and glide coda do not agree in the [round] feature.
(8) Agree[bk]-GV

Assign one violation mark for every syllable whose vowel and prenuclear glide do not agree in the [back] feature.
(9) Agree[rd]-GV

Assign one violation mark for every syllable whose vowel and prenuclear glide do not agree in the [round] feature.
(10) Ident[F]-G

Assign one violation mark for every glide which has a different value for the $[\mathrm{F}]$ feature from its input, where $[\mathrm{F}]=[$ back $]$, [round $]$ or [nasal].

## (11) Ident[F]-V

Assign one violation mark for every vowel which has a different value for the [F] feature from its input, where $[\mathrm{F}]=[$ back $]$ or [round $]$.

Since the [back] feature and the [round] feature should be agreed in together, constraints (6) and (7) are ranked the same high, and so are constraints (8) and (9). Constraint (5) is never violated, so it is ranked the highest. Agree $[F]-\mathrm{VG}^{8}$ is always prior to $A_{\text {gree }}[F]-G V$, so (6) and (7) are ranked higher than (8) and (9). In order to satisfy the constraints above, the vowel, instead of the glide, changes its value, so (10) is ranked higher than (11). Tableaux $(12-14)$ illustrate how these constraints work. The parenthesized exclamation marks indicate that either violation is fatal. The Hasse Diagram is illustrated in (15).
(12) / чә/ $\rightarrow$ [че]

| /чә/ | Phonotactics | Agree[F]-VG |  | Agree[F]-GV |  | Ident[F]-V |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | [bk] | [rd] | [bk] | [rd] |  |
| a. $\rightarrow$ qe |  |  |  |  | * | * |
| b. чә |  |  |  | (*!) | (*!) |  |
| c. $\Psi \varnothing$ | *! |  |  |  |  | * |

[^5](13) /jow/ $\rightarrow$ [jow]

| /jəw/ | Phonotactics | Agree[F]-VG |  | Agree[F]-GV |  | Ident[F]-V |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | [rd] | $[\mathrm{bk}]$ | $[\mathrm{rd}]$ |  |  |
| a. $\rightarrow$ jow |  |  |  | $*$ | $*$ | $*$ |
| b. jəw |  |  | $*!$ | $*$ |  |  |
| c. jew |  | $(*!)$ | $(*!)$ |  |  | $*$ |

(14) /aw/ $\rightarrow$ [aw]

| /aw/ | Ident[F]-G |  | Agree[F]-VG |  | Agree[F]-GV |  | Ident[F]-V |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | [bk] | [rd] | [bk] |  |  |
| a. $\rightarrow \mathrm{aw}$ |  |  |  | * |  |  | * |
| b. aw |  |  | (*!) | (*!) |  |  |  |
| c. pw |  | *! |  |  |  |  | * |
| d. aj | *! |  |  |  |  |  |  |

(15) The Hasse Diagram of mid/low vowel assimilation

Рhonotactics


Example (12) shows how Agree[bk]-GV and Agree[rd]-GV both assign violation marks to rule out the candidate *[чə]. Example (13) proves that AgREE[F]-VG is ranked higher than Agree[F]-GV, for candidate (13c) will be wrongly predicted as the winner if Agree[F]-GV is ranked higher. Example (14) proves that Phonotactics should be ranked higher than Agree[F]-VG because it will be a problem to determine whether [aw] or *[pw] is the optimal candidate if these constraints are ranked the same high. Note that the constraint Ident[F]-G is only illustrated in example (14) because satisfying this constraint will lead to $[\Psi]$ in examples (12) and (13), which also violates the highranked constraint Phonotactics. For simplicity, Ident[F]-G is only listed when it is active.

These constraints and their ranking successfully predict the correct output from each "phonemic group" (see 3.3). However, there are still some gaps that "should be legal" throughout this constraint ranking. For example, given an input /wow/, an optimal output *[wow] should be predicted, but it is not attested in Taiwanese Mandarin, as is illustrated in (16). Another problem lies in the rhymes with a high vowel following a prenuclear glide. For example, given an input /ju/, this constraint ranking will predict its output as [ji], which may be interpreted as a lengthened [i]. However, it seems not to be the fact, as shown in (17). These two problems will be later discussed in sections 4.1.2 and 4.1.3.
(16) /wəw/ $\rightarrow$ *[wow]

| /wəw/ | Рhonotactics | Agree[F]-VG |  | Agree[F]-GV |  | Ident[F]-V |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | [bk] | [rd] | [bk] | [rd] |  |
| a. © wow |  |  |  |  |  | * |
| b. wəw |  |  | *! |  | * |  |

(17) /ju/ $\rightarrow$ ? $[\mathrm{ji}]$

| /ju/ | Phonotactics | Agree[F]-VG |  | Agree[F]-GV |  | Ident[F]-V |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $[\mathrm{bk}]$ | $[\mathrm{rd}]$ | $[\mathrm{bk}]$ | $[\mathrm{rd}]$ |  |
| a. $\boldsymbol{\sigma}^{*} \mathrm{ji}$ |  |  |  |  |  | $*$ |
| b. ju |  |  |  | $(*!)$ | $(*!)$ |  |
| c. jy |  |  |  |  | $*!$ | $*$ |

### 4.1.2 The Phonotactic Restriction of Triphthongs

As what is mentioned in 4.1.1, rhymes such as *[jej], *[jaj] (see 3.2.3), *[wow], *[waw], *[чej], *[чај], *[чоw] and *[чaw] should be surfaced out because all of these patterns are actually the optimal choices throughout the constraint ranking. It seems that there are some more constraints which are active to these triphthongs.

To observe these absent patterns, it is found that glide [j] cannot precede a diphthong with a coda [j], glide [w] cannot precede a diphthong with a coda [w], and glide [ Y ] cannot precede any diphthongs, as is simplified in (18) below.
(18) Triphthongs that are not allowed
a. $\quad * \mathrm{j} \underline{\mathrm{V}}$
b. ${ }^{*} \mathrm{w} \underline{\mathrm{V}}$
c. ${ }^{*} \mathrm{q} \mathrm{V} \mathrm{j}$
d. ${ }^{*} \underline{Y} \underline{w}$

Regarding to this phenomenon, Duanmu (2007) has also proposed an analysis. In his opinion, there is a "dissimilation effect" interfering between the prenuclear glide
and the postnuclear glide. In Taiwanese Mandarin (as well as other dialects), the prenuclear glide and the postnuclear glide should have different values for both of their [back] and [round] features. This is the reason why there are no such rhymes in Taiwanese Mandarin where the prenuclear glide and the postnuclear glide are identical, and $[\mathrm{u}]$ cannot precede any diphthongs, for it is [-back] as [j] and [+round] as [w].

To deal with this dissimilation phenomenon, the Obligatory Contour Principle (henceforth OCP) is usually used. OCP was first developed by Leben (1973) for tonal dissimilation in African languages. The definition is shown in (19) below.

## (19) The Obligatory Contour Principle (Ito and Mester, 1998)

Adjacent identical elements are not allowed at the melodic level.

As (19) shows, OCP can only confine that two adjacent elements should not be identical (at least in the value of a certain feature). However, the problem going to be coped with here does not lie in an adjacent pairs, but two glides with a vowel in between. To solve this problem, I refer to Ito and Mester's (1998) interpretation to OCP. In their opinion, OCP is in fact a self-conjunction effect.

Take a look at Hsiao's (2000) example about tone sandhi in Southern Min. In Southern Min, only the last syllable in a tone sandhi domain remains its base tone; other syllables will undergo tone sandhi. This phenomenon seems to be able to be predicted with an OCP constraint as (20) shows, where T stands for base tone. However, this is awkward because OCP is used to prevent two adjacent segments from having identical "tone level" or "tone contour", while the base tones for each syllable are not the same.
(20) Ocp-T

Assign one violation mark for every adjacent pair of segments which remain the base tone.

Nevertheless, if we adopt Ito and Mester's interpretation to revise this OCP constraint as a self-conjunction, as (21) shows, the original definition of OCP will not be violated, and the constraint still remains its essence.
(21) $* \mathrm{~T}^{2}{ }_{\delta}$

Assign one violation mark for every tone sandhi domain where there are two syllables of the base tone.

Adopting this theory, the dissimilation phenomenon of the two glides in a same syllable can be enforced by a self-conjunction. This way, we can rule out the ill-formed outputs without violating the original definition of OCP.

Reviewing the unallowed triphthongs listed in (18) and Duanmu's analysis that the prenuclear and postnuclear glides should not share the same value for either [back] or [round] feature, we can introduce a self-conjunction to prohibit a syllable from having two glides that have the same value for the [back]/[round] feature, as is in (23), which is self-conjoined by constraint (22) below.
(22) * $\mathrm{G}_{\mathrm{aF}}$

Assign one violation mark for every glide of $[\alpha \mathrm{F}]$ feature, where $\mathrm{F}=[$ back $]$ or [round].

## (23) $*\left[\mathrm{G}^{2}{ }_{\alpha \mathrm{F}}\right]_{\sigma}$

Assign one violation mark for every syllable which contains two glides of $[\alpha \mathrm{F}]$ feature, where $\mathrm{F}=[$ back $]$ or [round].

Constraint (23) successfully rules out all the ill-formed triphthongs. However, when these candidates are ruled out, there should be another candidate to be selected as an optimal output. To solve this problem, here we have to mention the experiment in 3.2.3 again. The result of the experiment shows that Taiwanese Mandarin speakers tend to pronounce the word yai 'cliff' [jaj] as [aj] instead of [ja], which indicates that it is the prenuclear glide that is dropped when there are two glides having the same value for [back] or [round] feature in a syllable. The descriptive generalization is shown in (24). In addition to constraints (22) and (23) introduced above, there are still other constraints active, as below. Constraint (25) indicates the strategy of deleting the prenuclear glide. Constraint (26) is also listed to show another possible strategy to repair the ill-formed triphthongs. In addition to the strategy of deleting the coda, changing the coda might also be a method to satisfy the higher-ranked constraints. Therefore, $\operatorname{Ident}[F]-G$ is also active here.
(24) The descriptive generalization for the phonotactic restriction of triphthongs

The prenuclear glide and the postnuclear glide of a syllable should have different values for both [back] and [round] feature. This requirement is enforced by prenuclear glide deletion.
(25) Max

Assign one violation mark for every segment in the input which does not have a correspondent output.

Assign one violation mark for every coda in the input which does not have a correspondent output.

Constraints (25) and (26) both assign a violation mark for deletion; however, (26) has a stricter definition, which is only violated when a coda is deleted, while (25) is violated no matter which segment does not surface out. Logically, it is known that when a candidate violates constraint (26), it must violate (25) as well; but not vice versa. Therefore, it is certain that (26) should be absolutely ranked higher than (25).

To observe the result of the [jaj]-to-[aj] experiment, it is observed that there should not be two [j]'s in the syllable, so $*\left[\mathrm{G}^{2}{ }_{\alpha \mathrm{F}}\right]_{\sigma}$ is satisfied, and it absolutely dominates $* \mathrm{G}_{\alpha \mathrm{F}}$, for it is a self-conjunction (see. 2.1.3). In fact, MAx should dominate $* \mathrm{G}_{a \mathrm{~F}}$ because $* \mathrm{G}_{a \mathrm{~F}}$ is violated by every candidate containing a glide. Therefore, MAx should be ranked higher than $* \mathrm{G}_{\alpha \mathrm{F}}$ to avoid glide deletion. The prenuclear glide [j] is deleted, so MAX is violated, but the coda [j] is not deleted, so Max-Coda is satisfied. Both of the glides do not change any features, so Ident[F]-G is satisfied. Through these descriptions, we can rank the satisfied constraints higher and the violated ones lower, as tableau (27) shows.
(27) /jaj/ $\rightarrow$ [aj]

| /jaj/ | * $\left[\mathrm{G}^{2}{ }_{\alpha \mathrm{F}}\right]_{\sigma}$ | Ident[F]-G | Max-Coda | Max | $* \mathrm{G}_{\alpha \mathrm{F}}$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| a. $\rightarrow$ aj |  |  |  | * | * |
| b. jaj | *! |  |  |  | ** |
| c. ja |  |  | *! | * | * |
| d. waj |  | *! |  |  | ** |
| e. jaw |  | *! |  |  | ** |
| f. a |  |  | *! | ** |  |

Note that candidate (e) in tableau (27) changes its vowel from the input. Here we do not have a candidate *[jaw], which violates Ident[F]-G as well but remain the same vowel as its input because *[jaw] is absolutely a worse choice than [jaw] no matter how these constraints are ranked, for we have already discussed the low vowel assimilation phenomenon in 4.1.1. The Hasse Diagram of the triphthong phonotactic restriction is illustrated in (28).
(28) The Hasse Diagram of the phonotactic restriction of triphthongs


As (28) shows, the example in tableau (27) shows that the optimal output for /jaj/ is [aj], which violates MAX (and of course ${ }^{*} \mathrm{G}_{\alpha \mathrm{F}}$ ). This indicates that the violations of
*[ $\left.\mathrm{G}^{2}{ }_{a \mathrm{~F}}\right]$, IDEnt[F]-G and Max-Coda are not allowed. Therefore, all of these constraints dominate Max. Furthermore, since glides should not be deleted in most contexts (for example, $/ \mathrm{ja}$ / does not surface as [a]), MAX dominates ${ }^{*} \mathrm{G}_{\mathrm{aF}}$.

The constraint ranking developed in this section may also explain the absence of the high vocoid sequences $*[\varphi i]$ and $*[\varphi u]$, for the two high vocoids which have the same values for [back] or [round] features are in the same syllable no matter whether there is a mid/low vowel between them or not. However, this analysis is controversial because the constraint $*\left[\mathrm{G}^{2}{ }_{\alpha \mathrm{F}}\right]_{\sigma}$ only prohibits two glides that have the same values for [back] or [round] features from appearing in the same syllable, while [i] and [u] are actually vowels. If we transcribe the sequence as two glides, it is still awkward for there will be no peak in the syllable. Moreover, in addition to *[чi] and *[чu], *[ju] and *[wi], which are also sequences of two high vocoids, are also not allowed to surface. The absence of these patterns is going to be discussed in the next section.

### 4.1.3 The Sequence of High Vocoids

As the previous section mentions, here we are going to discuss patterns composed of two high vocoids, none of which is surfaced out. Duanmu (2007) considers *[ju] and *[wi] to be accidental gaps, but it is apparent that the absence of these patterns are systematic. Regarding to this phenomenon, Lin (2007) has mentioned that "glides cannot be followed by a high vowel which has a different value for the [back] feature." This statement here does not rule out the ill-formed pattern *[yi]. In fact, it is weird for Lin to specify "what kind of high vowel cannot follow a glide" because there is actually no high vowel following a glide.

Via revising Lin's statement, we can generalize that two high vocoids cannot be adjacent, or glides cannot be followed by a high vowel. Since it is two adjacent 40
segments that are prohibited to share a certain feature, OCP, which is mentioned in 4.1.2, can be introduced as a constraint here, as in (29).
(29) Ocr-HIVocoid

Assign one violation mark for every pair of adjacent high vocoids.

In fact, it is not necessary to specifically prohibit two adjacent "high vocoids" based on the analysis in this section. Instead, we can just prohibit two adjacent segments that are both [+high]. However, if we substitute a more general constraint called "OcpH!" for constraint (29), it will wrongly predict the output of/in/ as *[en], which will be elaborated in 4.2.1.2. Therefore, in order to take the analyses of the following sections into account, here we have to specifically define that it is two high vocoids that cannot be adjacent.

In the previous section, we predict that the ill-formed triphthong *[jaj] is repaired by deleting the prenuclear glide according to how Taiwanese Mandarin speakers pronounce the word yai 'cliff'. Likewise, now that the sequences of two high vocoids have been ruled out by constraint (29), there should be optimal outputs for these inputs. Then, how do the sequences of two high vocoids surface out? The answer can be found from the evidence of loanwords and translated names. The examples are shown in $(30-31)$ below.
(30) Loanwords and translated names containing the sequence of [ju] in the input

| English |  |  | Taiwanese Mandarin |  |
| :--- | :--- | :--- | :--- | :---: |
| a. UFO | [ju.for] | youfu | [jow.fu] |  |
| b. Ulysses | [ju.lı.siz] | Youlixisi | [jow.li.ci.sz] |  |
| c. New Zealand | [nju zi.lænd] | Niuxilan | [njow.ci.lan] |  |
| d. Hugh Jackman | [hju dzæk.mən] | Xiu Jiekeman | [cjow tcje.k ${ }^{\text {h }}$ ə.man] |  |

(31) Loanwords and translated names containing the sequence of $[\mathrm{wI}]^{9}$ in the input


Examples in (30) and (31) show that when a loanword or a translated name has a sequence of two high vocoids in its English origin, there will be a mid vowel inserted between the high vocoids in the translated form, and of course, the vowel will undergo a glide formation, for there should not be two vowels in a single syllable. The descriptive generalization is shown in (32).
(32) The descriptive generalization for the sequence of high vocoids
a. Two high vocoids cannot be adjacent. This requirement is enforced by mid vowel insertion.

[^6]b. There is only one vowel in a single syllable. This requirement is enforced by glide formation of a high vowel.

According to (32a), we developed constraint (33) as a violated faithfulness constraint for an output satisfying Ocp-HIVocoid. Constraint (34) is developed to ensure the inserted vowel is a mid one instead of a low one. Constraints (35-36) are developed based on (32b). Constraint (35) is based on McCarthy's (2008) list of common markedness constraint, which includes constraints called "Comp-Onset" and "СомpCODA" assigning violation marks for consonant cluster in the onset or coda position. Here the candidates containing a vowel cluster should be ruled out, so a new constraint "Comp-Nuc" is introduced, where Nuc stands for nucleus. Constraint (36) is used for vowel gliding since a vowel changes its value for the [syllabic] feature when it becomes a glide. In addition to the constraints mentioned in this paragraph, MAX is also active here. Tableau (37) provides an example of how a sequence of two high vocoids surfaces out.
(33) Dep-MidV

Assign one violation mark for every mid vowel in the output which does not have a correspondent input.
(34) Dep-LowV

Assign one violation mark for every low vowel in the output which does not have a correspondent input.
(35) *Comp-Nuc

Assign one violation mark for every vowel cluster in the nucleus position.
(36) Ident[syl]

Assign one violation mark for every segment which has a different value for the [syllabic] feature from its input.
(37) $/ \mathrm{ju} / \rightarrow[$ jow $]$

| /ju/ | Ocp- <br> HIVocoid | Dep-LowV | Max | Dep-MidV | * Comp- <br> Nuc | Ident[syl] |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| a. $\rightarrow$ jow |  |  |  | * |  | * |
| b. jaw |  |  | - |  |  | * |
| c. u |  |  | *! |  |  |  |
| d. ju |  |  |  |  |  |  |
| e. jou |  |  |  |  | *! |  |
| Diff |  |  |  |  |  |  | indicates that the constraints left to it and right to it have no domination relations. Let's first take a look at candidates $(37 \mathrm{a}-\mathrm{d})$. The optimal candidate is $(37 \mathrm{a})$ [jow], which has an inserted mid vowel. Other candidates worse than (37a) include (37b) [jaw], which has an inserted low vowel, (37c) [u], which deletes the prenuclear glide, and (37d) *[ju], which has a sequence of two high vocoids. Among these candidates, the one violates Dep-MidV is the most tolerable. Therefore, we can have a result that OcpHiVocoid, Dep-LowV and Max all dominate Dep-MidV.

On the other hand, if we compare candidates (37a) and (37e), the violation differs in the fact that the former violates Ident[syl] while the latter violates *Comp-Nuc. The better output is (37a) [jow], which proves that *Comp-Nuc dominates Ident[syl]. Here I have to clarify that the transcription of (37e) [jou] is actually widely used in previous
studies. In this thesis, however, it is necessary to view a diphthong as a vowel plus a glide coda because it has been discussed that the prenuclear glide and the postnuclear glide should not have the same value for their [back] and [round] features (see 4.1.2). If we transcribe the postnuclear coda as a vowel, it will be considered as a part of the nucleus, and then the constraint prohibiting the prenuclear glide and the postnuclear glide to have the same value for their [back] and [round] features will become weird, for it is therefore regulating a "segment" and a "part of segment" to share the same value for the features. As a result, in order not to be contradicted, here we introduce these constraints to ensure the vowel gliding rule.

Now two constraint rankings have been developed, as the doubled frame line in tableau (37) indicates. The fact is that we have no evidence to prove that there are domination relations between the constraints left to the doubled frame line and the ones right to it. To illustrate the constraint ranking clearer, the Hasse Diagram is available in (38). It does not matter for the candidates to be evaluated by the ranking on the left side or the right side first because [jow], the optimal output is always the last one to be ruled out.
(38) The Hasse Diagram of the sequence of high vocoids


As (38) shows, to repair a rhyme composed of two adjacent high vocoids, the best policy is to insert a mid vowel to violate Dep-MidV. Dep-MidV is thus dominated by $^{\text {b }}$ Ocp-HiVocoid, Dep-LowV and Max. On the other hand, in order to make sure that a
high vowel becomes a glide when there is another vowel, *Comp-Nuc should dominate Ident[syl].

### 4.1.4 Summary

We have discussed all the rhyme patterns that do not contain a nasal coda in Taiwanese Mandarin through the analyses so far, and have obtained the conclusions as below:
(i) Mid vowels and low vowels agree with the glide coda in [back] and [round] feature. (see 4.1.1)
(ii) Mid vowels and low vowels agree with the prenuclear glide in [back] and [round] feature when there is no glide coda. (see 4.1.1)
(iii) When two glides that share the same values for [back] or [round] feature are in the same syllable, the prenuclear one fail to surface out. (see 4.1.2)
(iv) A mid vowel is inserted between two adjacent high vocoids. (see 4.1.3)

These results can almost explain the presence and the absence of the rhymes without a nasal coda listed in 3.1. As for the constraint ranking, we have developed the Hasse Diagrams at the end of each subsection. However, the constraint rankings should not be independent. After all, they are all evaluating the patterns of the same language, and even the same dialect. In addition, the constraints Max and Ident[F]-G are active in not only one analysis. Therefore, it is possible to combine the Hasse Diagrams in each subsection. Though not all of them have domination relations with another, the system is clearer if we have all the active constraints in one Hasse Diagram, as is illustrated in (39) below.
(39) The Hasse Diagram of rhymes without a nasal coda


To have an overview of all the rhyme patterns without a nasal coda, the constraint ranking in diagram (39) can correctly predict their presence and absence, except for an unallowed rhyme *[чa]. Because there are no other absent patterns that seem to be symmetric to this rhyme, it is reasonable to consider the absence of *[ya] an accidental gap.

### 4.2 Rhymes with a Nasal Coda

### 4.2.1 High Vowel Preceding a Nasal Coda

In section 2.2.1, it is mentioned that $\operatorname{Lin}(2007)$ has a hypothesis stating that there are two forms of the $\underline{\mathbf{Y X}}$ structure, diphthongs and "high/low vowel + nasal." Based on her assumption, mid vowels never precede nasal codas in the underlying representation. In other words, all the mid vowels preceding a nasal coda in the surface representations are inserted or derived from other underlying forms. Her theory asserts that an [ 2 ] is 47
inserted between a high vowel and a nasal coda when they have different values for the [back] feature. This indicates that the structure of "high vowel + nasal" and that of "glide $+\partial+$ nasal" belong to the same phonemic group (see 3.3). A brief chart is arranged in (40) below.
(40) The outputs of underlying "high vowel + nasal"

| a. in | *jən | b. | *un | wən | c. | yn | *чən |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| d. in | *jəŋ | e. | *un | *wəŋ | f. | *yŋ | *чəŋ |

There are two problems having to be solved in chart (40). First, there should be accurately one allowed rhyme in each item, but no rhymes can be surfaced out in (e) and (f). The other problem lies in (d). According to Lin's theory, an [ə] should be inserted when a high vowel and its adjacent nasal coda have different values for the [back] feature, but it is [in] that is surfaced rather than *[joy], with [i] and [ y ] differing in the [back] feature.

### 4.2.1.1 The Regular Patterns [in], [won] and [yn]

Compare the patterns in $(40 \mathrm{a}-\mathrm{c})$ to those in $(40 \mathrm{~d}-\mathrm{f})$, it is known that patterns with an alveolar nasal $[\mathrm{n}]$ are more regular than those with a velar nasal $[\mathrm{n}]$. Therefore, the patterns with [ n ] is going to be analyzed first. As Lin's (2007) hypothesis, an [ə] is inserted between only when the high vowel and its following nasal coda have different values for the [back] feature. Based on this phenomenon, a preliminary descriptive generalization is introduced in (41), and the related constraints are illustrated in (42 47) below. Note that it is necessary to specify the vowel as "high" ones because [u] and
[ə] are both [+back]. If we assume that all vowels should agree with its nasal coda in [back] feature, then inserting an [ə] between [ u ] and $[\mathrm{n}]$ does not help repair the ill form.
(41) A preliminary descriptive generalization for high vowel preceding a nasal coda A high vowel should have a different value for its [back] feature from its nasal coda. This requirement is enforced by [ə]-insertion.
(42) Agree[bk]-V $\mathrm{V}_{\mathrm{H}} \mathrm{N}$

Assign one violation mark for every syllable whose high vowel and nasal coda do not agree in the [back] feature.
(43) Dep-ə

Assign one violation mark for every [ə] in the output which does not have a correspondent input.
(44) Dep-e

Assign one violation mark for every [e] in the output which does not have a correspondent input.
(45) Dep-o

Assign one violation mark for every [ o ] in the output which does not have a correspondent input.
(46) Ident $[\mathrm{bk}]-\mathrm{N}$

Assign one violation mark for every nasal coda which has a different value for the [back] feature from its input.

Assign one violation mark for every high vowel which has a different value for the $[\mathrm{F}]$ feature from its input, where $[\mathrm{F}]=[$ back $]$ or [round $]$.

Constraint (42) avoids a high vowel to have a different value for its [back] feature from the following nasal coda. Constraint (43) indicates the strategy of inserting an [ə], which is violable, so it is ranked low. Constraints (44) and (45) are developed to avoid inserting other mid vowels rather than [ə]. Agree[F]-GV (which in fact includes two constraints) acts as a low-ranked constraint here, for [ə] does not agree in [back] and [round] features at the same time with any glide. Constraints $(46-47)$ are here to prevent the vowel and the nasal from changing their features. Tableau (48) is the example of how these constraints are ranked.
(48) /un/ $\rightarrow$ [wən]


In 4.1.1, it is proved that Agree[F]-GV dominates Ident[F]-V. For example, /jz/ is surfaced as [je] to satisfy Agree[F]-GV by violating Ident[F]-V. However, in tableau (48), candidate (d), which violates the lower-ranked $\operatorname{Ident}[F]-V$, is worse than candidate (a), which violates the higher-ranked $\left.\operatorname{Agree}^{[r d]}\right]$ GV. Therefore, it is assumed that there should be a constraint that is ranked even higher to rule out (d) [yn]. For this reason, we rank the constraint $\operatorname{Ident}[F]-\mathrm{HIV}^{2}$ higher than Agree[F]-GV, which specifically $^{\text {a }}$ restricts a high vowel to change its feature. This is very reasonable because there are three high vowel phonemes in Mandarin but only one mid/low phoneme. Changing the features of a high vowel may lead to confusion.

There are still some facts worth mentioning. First, some may argue that there are several possible candidates not listed in tableau (48), such as *[uən], [wan], [u] and *[n]. In section 4.1.3, we have developed a constraint * Comp-Nuc to avoid vowel clusters, so *[uən] is absolutely worse than [wən], for *Comp-Nuc >> Ident[syl]. As for [wan], also mentioned in 4.1.3, inserting a low vowel is worse than inserting a mid vowel, for DepLowV >> Dep-MidV, so [wan] is also absolutely worse than [wən]. Candidates [u] and [ $n$ ] violate Max, which has been proved ranked higher than $\mathrm{Dep}_{\mathrm{E}}-\mathrm{Mid} \mathrm{V}$, so there are no reasons for them to be the optimal outputs.

Second, though [o] does not agree with [ n ] in [back], it does agree with the prenuclear glide [w]. Therefore, Dep-o still needs to be developed here and ranked higher than Agree[F]-GV. Furthermore, in 4.1.3, a constraint Dep-MidV is developed to assign violation marks for every mid vowel which does not have a correspondent input. The constraint is there to ensure that a mid vowel is inserted between two high vocoids. (e.g. $/ \mathrm{ju} / \rightarrow$ [jow].) In this section, it is found that inserting an [ə] is better than inserting an [e] or [o], even though all of them are mid vowels. Therefore, there should be stricter constraints to define inserting which mid vowel is better or worse, just like what Max and Max-Coda in 4.1.2 do. The example of /ju/ becoming [jow] inserts an [o], which
therefore violates not only Dep-MidV, but also the stricter Dep-o, as (49) shows. Likewise, when it comes to the example of /wi/ $\rightarrow$ [wej], the constraint ranking works the same. Note that Dep-e is not active in example (49) because *[jew] is absolutely a worse choice than [jow], for it violates Agree[F]-VG. Neither is Dep-o active in (50). Just make sure that Dep-e and Dep-o are always ranked together.
(49) $/ \mathrm{ju} / \rightarrow[$ jow $]$


| /wi/ | Ocp-HiVocoid | Dep-LowV | Max | Agree[F]-VG | Dep-e | Dep-ə |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| a. $\rightarrow$ wej |  | heng | nl |  | $*$ |  |
| b. waj |  | $*!$ |  |  |  |  |
| c. i |  |  | $*!$ |  |  |  |
| d. wi | $*!$ |  |  |  |  |  |
| e. wәj |  |  |  | $*!$ |  | $*$ |

By reviewing the mid vowel insertion between two high vocoids, it is found that there is a different between these two phenomena: The sequence of two adjacent high vocoids is repaired by inserting [e]/[o] while the sequence of VN with different values
for the [back] features is repaired by inserting [ə]. Tableaux (49) and (50) have helped explain the difference. The choice of [wən] instead of *[wen] proves that Dep-e (as well as Dep-o) dominates Dep-a, but [wej] is better than *[wəj], for Agree[F]-VG is ranked even higher than Dep-e.

The constraint ranking is summarized as Hasse Diagrams in (51). Diagram (a) illustrated the constraint ranking for the [ə]-insertion between high vowel and nasal; (b) revises the constraint ranking for the $[\mathrm{e}] /[\mathrm{o}]$-insertion between two high vocoids in 4.1.3; (c) concludes (a) and (b) by combining them together.
(51) The Hasse Diagrams of mid vowel insertion
a. [ə]-insertion between high vowel and nasal

b. $[\mathrm{e}] /[\mathrm{o}]$-insertion between two high vocoids


Dep-ə
c. Mid vowel insertion


In (51a), Dep-ə is ranked the lowest because inserting an [ə] is the best policy to repair the ill forms. At the same time, since Agree[bk]-GV and Agree[rd]-GV are both violated, for [ə] does not agree with the prenuclear glide, they are also dominated by Dep-e, Dep-o and Ident[F]-HiV. On the other hand, Agree[bk]-VhN and Ident[bk]-N are only proved to dominate Dep-ə because they do not contradict with Agree[bk]-GV and Agree[rd]-GV.

Hasse Diagram (51b), revised from the Hasse Diagram (38) in 4.1.3, separates the constraint Dep-MidV into Dep-e, Dep-o and Dep-2, with Dep-e and Dep-o dominating Dep-a. In these cases, the mid yowels inserted between the high vocoids are [e] and [o] instead of [ə] in order to satisfy Agree[bk]-VG and Agree[rd]-VG, so they also dominate Dep-e and Dep-o. Combining (51a) and (51b) can obtain diagram (51c).

### 4.2.1.2 The Irregular Outputs of /ing/, /uy/ and /yy/

Based on the constraint ranking in (51), Agree[bk]- $\mathrm{V}_{\mathrm{H}} \mathrm{N}$ is a high-ranked constraint which is inviolable. However, the rhyme [ig] does violate this constraint. The output of $/ \mathrm{iy} /$ is predicted as *$[j \partial y]$ through this constraint ranking, as (52) shows, but *[jəy] does not appear in Taiwanese Mandarin. Here we do not consider other candidates like [joy] or *[jey] because such forms are proved worse in the previous part.
(52) $/ \mathrm{ig} / \rightarrow$ *[jəy]

| /iy $/$ | Agree[bk]-VHN $^{\text {Dep-2 }}$ |  |
| :--- | :---: | :---: |
| a. © jəy |  | $*$ |
| b. in | $*!$ |  |

Because the constraint ranking wrongly predicts the output of /iy/, it is assumed that there is a constraint ranked even higher to rule out the predicted output *[jəy]. To further observe the patterns, it is found that in addition to the absence of *[jəŋ], *[чəŋ], which is predicted as the output of $/ \mathrm{yy} /$, is also not allowed. Hence, we assume that the combination of [əŋ] is very marked in Taiwanese Mandarin, which blocks *[jəŋ] and *[чəŋ] from surfacing out. The descriptive generalization is thus revised as (53), and the constraint is shown in (54) with its definition.
(53) The revised descriptive generalization for high vowel preceding a nasal coda When a high vowel has a different value for the [back] feature from its nasal coda, an [ə] is inserted, unless it leads to the sequence of [əŋ].

Assign one violation mark for every sequence of [əŋ].

This constraint seems ad hoc. However, Taiwanese Mandarin is not the only dialect in which the sequence of $[\partial \eta]$ has something to do with irregular patterns. In Beijing dialect, the word weng 'old man' is pronounced as [wəy] (Lin, 2007), where an [ə] is inserted even though the high vocoid and the nasal coda are both [+back]. This example supports that [əๆ] is exactly a highly-marked sequence. Because [iŋ] is a better candidate than $*[j ə \eta]$, the constraint $* \partial \eta$ should dominate Agree $\left.^{2} \mathrm{bk}\right]-\mathrm{V}_{\mathrm{H}} \mathrm{N}$, as is illustrated in tableau (55).
(55) $/ \mathrm{in} / \rightarrow[\mathrm{iy}]$

| /in $/$ | *วy | AGReE[bk]-VHN | DEP-ə |
| :--- | :---: | :---: | :---: |
| a. $\rightarrow$ iŋ |  | $*$ |  |
| b. jəŋ | $*!$ |  | $*$ |

The constraint *әy successfully rules out *[jəŋ] as well as *[чəŋ], a possible output of $/ \mathrm{yy} /$, but $*[\mathrm{yn}]$ is also not surfaced out even though it should be the optimal output according to the constraint ranking in (55). Now that neither *[yŋ] nor *[чəŋ] is allowed, then how does /yy/ surface out? According to the "phonemic group" introduced in 3.3, when $\underline{\mathbf{X}}=[\mathrm{n}]$ or $[\mathrm{n}]$, there should be only one choice among $\emptyset,[\mathrm{e}],[\rho]$ and $[\mathrm{o}]$, as the shadowed boxes in chart (56) illustrate. Nevertheless, the list in (57) below proves that there is something wrong under this assumption.
(56) Phonemic group indication chart $(\underline{\mathbf{G}}=$ any $)$

| $\underline{\mathbf{V}} \backslash \underline{\mathbf{x}}$ | $\emptyset$ | $\mathrm{\emptyset}$ | w | n | n |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $(\mathrm{i}, \mathrm{u}, \mathrm{y})$ |  |  |  |  |  |
| e |  | $C h e n$ | $h i$ |  |  |
| $\partial$ |  |  |  |  |  |
| o |  |  |  |  |  |
| a |  |  |  |  |  |
| a |  |  |  |  |  |

（57）The combinations of $\underline{\mathbf{G V N}}$ ，where $\underline{\mathbf{V}}=\emptyset$ ，e，$\partial$, o．The arrows illustrate the derivation prediction．

| a．＊n | ＊en | ən | ＊ on |  | ＊un | ＊wen | wən | ＊Won |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| b．＊$\quad$ y | ＊eŋ | əリ | OI |  | *uy | ＊wey | ＊Wəり | ＊Woy |
| c．in | jen | ＊jən | ＊jon | g． | yn | yen | ＊$¢$ ¢ | ＊¢о高 |
| d．in | ＊jey | ＊jəŋ | jon | h． | *yy | ＊чеу | ＊чəŋ | ＊¢о才 |

The underlined gray patterns in $(57 \mathrm{c})$ and $(57 \mathrm{~g})$ will be later discussed in 4．2．3．2， for the underlying form of the vowels are low ones．In addition to these patterns，it is found that there are two allowed patterns in（57b）and（57d），while none is allowed in （57f）and（57h）．Because the rhyme［joy］in（57d）contains both［－back］and［＋round］ features of［y］，it is assumed that $/ \mathrm{yy} /$ is surfaced as［jon］，and symmetrically，／uy／is surfaced as［oŋ］．

The fact that $/ \mathrm{yy} /$ is surfaced as［jon］indicates that the vowel $/ \mathrm{y} /$ in the input undergoes a process of splitting．Historically，［y］is the combination of［ju］，with［j］ reflecting the［－back］feature and $[\mathrm{u}]$ the［＋round］feature，and all of the three being ［＋high］（Duanmu，2007）．Splitting the［y］preceding［y］into two segments does prevent a［－back］vowel from being adjacent to a［＋back］nasal coda and avoid the sequence of $[ə \eta]$ ，so this is the best strategy to repair the ill－formed $*[y \eta]$ ．Splitting a segment violates a faithfulness constraint，though．Constraint（58）is therefore developed； however，tableau（59）still wrongly predicts the output of $/ \mathrm{yy} /$ as＊［juy］，for［y］is split into［ju］instead of［jo］．
(58) Integ

Assign one violation mark for every segment in the input which has more than one correspondent outputs.
(59) $/ \mathrm{yy} / \rightarrow$ *[juy]

| $/ \mathrm{yy} /$ | *วり | AGree[bk]-V HN | DEP-ə | InTEG |
| :--- | :---: | :---: | :---: | :---: |
| a. juy |  |  |  | $*$ |
| b. yy |  | $*!$ |  |  |
| c. чəๆ | $*!$ |  | $*$ |  |

There are two explanations for the ill-formed *[juy]. First, as discussed in 4.1.3, two high vocoids cannot be adjacent. Splitting [y] into [ju] does violate the constraint Ocp-HiVocoid. The strategy to repair the ill forms is to insert [e] or [ o ] between the high vocoids. By doing so, other ill-formed syllables are derived, such as *[jouy] and *[jown]. To rule out *[jouy], the constraint *Comp-Nuc should be active here, but gliding the high vowel only leads to another ill form *[jowy], which should be ruled out by introducing a new constraint to prohibit complex codas. The only strategy to avoid all these ill forms is to change the value for the [high] feature of the vowel, and the optimal output [jon] can finally be derived. This analysis surely explains why [jon] is chosen as an optimal output for $/ \mathrm{yn} /$, but there is another explanation, which is much more economic.

To make the other explanation of why $/ \mathrm{yy} /$ is surfaced as [joy], we have to focus on the output of /uy/ first. As the assumption of $/ \mathrm{yy} /$ surfacing as [joy], it is found that both [əy] and [oy] in (57b) are allowed while no patterns can be surfaced in (57f). Because of the [+back] and [+round] features retained in [o], it is assumed that [oy] is
derived from /un/. Now that/un/ is surfaced as [on], it is not surprised for *[juy] to be repaired as [joy]. Therefore, by explaining the derivation of [oy], we can at the same time solve the problem of the surface form of $/ \mathrm{yy} /$.

Actually, the phenomenon that $[\mathrm{u}]$ is lowered as [ o ] when preceding a velar coda is also observed in other Sinitic languages, such as Southern Min and Cantonese. In Southern Min, there are three back vowels, namely [o], [0] and [u]; however, only [o] can precede a velar coda [ n$]$ or [ k$]$, which means that the combinations of a high back vowel and a velar coda *[uy] and *[uk] are prohibited in Southern Min (Chung, 1995). On the other hand, in Cantonese, there is a distinction between the rhymes [ 9 y$]$ and $[\mathrm{o} \square$ ] and between [ok] and [ok], as the examples in (60). However, [o] does not solelyexist in the rime and moreover, there is no *[un] or *[uk]. Therefore, it is assumed that [on] and [ok] are also derived from/uy/ and /uk/ (Yip and Matthews, 2006).
(60) The minimal pairs of [ $\rho$ ] and [ o$]$ in Cantonese

| mok | $[\mathrm{mok}]$ | 'do not' | muk | $[\mathrm{mok}]$ | 'wood' |
| :--- | :--- | :--- | :--- | :--- | :--- |
| dong | $[\mathrm{ton}]$ | 'swing' | dung | $[\mathrm{toy}]$ | 'move' |

Although the stop coda $[\mathrm{k}]$ is not allowed in Mandarin, it is observed that in other languages where $[\mathrm{u}]$ is lowered when preceding the velar nasal $[\mathrm{n}]$, like Southern Min and Cantonese, the velar stop [k] also triggers the [u]-lowering. Therefore, we conclude that two high back segments cannot be adjacent in the rhyme. The descriptive generalization is illustrated in (61), and constraints $(62-63)$ are therefore introduced.
(61) The descriptive generalization for [u]-lowering

Two high back segments cannot be adjacent in the rhyme domain. This requirement is enforced by lowering the vowel.
(62) Ocp-[+hi, +bk]

Assign one violation mark for every pair of adjacent high back segments.
(63) Ident[hi]-V

Assign one violation mark for every vowel which has a different value for the [high] feature from its input.

In addition to constraints (62) and (63), *әy and Dep-o are also active here as highranked constraints to avoid predicting ill-formed candidates like *[wəy] and *[woy] as an output, as illustrated in tableau (64). Note that candidate (64c) also violates Dep-r, but this constraint is not considered here because Dep-o has been proved to dominate Dep-ə and the candidate should have been ruled out when evaluated by *əy. The Hasse Diagrams in (65) concludes the constraint rankings for the rhymes [iy], [oy] and [joy].
(64) $/ \mathrm{un} / \rightarrow[\mathrm{oy}]$

| $/ \mathrm{uy} /$ | OcP-[+hi, +bk] | *əy | DEP-o | Ident[hi]-V |
| :--- | :---: | :---: | :---: | :---: |
| a. $\rightarrow$ on |  |  |  | $*$ |
| b. un | $*!$ |  |  |  |
| c. wəy |  | $*!$ |  |  |
| d. woy |  |  | $*!$ |  |

(65) The Hasse Diagrams of rhymes underlyingly "high vowel $+[\eta]$ "
a. The unallowed sequence [əŋ]

b. [u]-lowering

c. Combining to predict [jon]


Diagram (65a) shows that * $\partial y$ is ranked the highest because it is inviolable. In
 when it comes to $/ \mathrm{yy} /$, Integ is violated to satisfy Agree $\left.^{\text {a }} \mathrm{bk}\right]-\mathrm{V}_{\mathrm{H}} \mathrm{N}$, so Agree $\left.^{\mathrm{l}} \mathrm{bk}\right]-\mathrm{V}_{\mathrm{H}} \mathrm{N}$ dominates Integ. That Agree[bk]- $\mathrm{V}_{\mathrm{H}} \mathrm{N} \gg$ Dep-ə has been illustrated in (51).

Diagram (65b) shows that when two high back segments are adjacent, the best policy to repair this ill form is lowering the vowel instead of inserting other vowels. Therefore, Ident[hi]-V is ranked the lowest. Combining (65a) and (65b) can obtain (65c), where DEP-o >> DEP-ə is proved in previous analysis.

### 4.2.2 [ən] and [əŋ] without a Prenuclear Glide

In 4.2.1.2, it is concluded that *[jəŋ] and *[чəŋ] do not surface out successfully as the outputs of $/ \mathrm{in} /$ and $/ \mathrm{yy} /$ because the marked sequence of [əŋ] is not allowed in Taiwanese Mandarin. However, there are a lot of examples in Taiwanese Mandarin which do have this sequence in the rhyme, such as deng [təy] 'light', keng [ $\left.\mathrm{k}^{\mathrm{h}} \partial \mathrm{\eta}\right]$ 'hole', sheng [səŋ] ‘voice, etc. Why are these syllables allowed while *[jəŋ] and *[чəŋ] are not?

To solve this problem, we have to mention the issue proposed in 2.2.1 again: Lin (2007) states that only high vowels and low vowels can precede a nasal coda in the underlying representation, and therefore the mid vowels preceding a nasal coda in the surface forms are inserted or derived from high/low vowels. Based on this hypothesis, we assume that the $[\partial]$ 's in the words like deng, keng, and sheng are also inserted. With this assumption, the underlying forms of the rhymes [ən] and [əๆ] are considered to be $/ \mathrm{n} /$ and $/ \mathrm{y} /$ respectively.

This hypothesis does not violate Lin's statement that there are no mid vowels preceding a nasal coda in the underlying representation, and it can also explain why the marked sequence [əŋ] can still surface out in syllables without a prenuclear glide (which have no vowels in the underlying forms), for Dep-ə is the lowest-ranked constraint among all of the constraints which prohibit a vowel to be inserted. To develop the constraints, a descriptive generalization is introduced first in (66), and constraint (67) is developed to ensure the existence of the nucleus. Ident[syl] is high-ranked here to avoid syllabic nasals. *วŋ is ranked lower than Nuc and Ident[syl] here, for the sequence of [əŋ] is surfaced in order to fulfill the requirement of having a nucleus. Though inserting an [ə] causes [əๆ], it is still a better choice than inserting an [o]. Therefore, Dep-o is also ranked high here. Tableau (68) shows how these constraints work.
(66) The descriptive generalization for inserting [ə] as a nucleus

A syllable should have a nucleus. This requirement is enforced by [ə]-insertion.
(67) Nuc

Assign one violation mark for every syllable which has no nucleus.
(68) $/ \mathrm{y} / \rightarrow[$ ŋ $]$

| /n/ | Nuc | Ident[syl] | Dep-o | * ${ }^{\text {y }}$ | Dep-ə |
| :---: | :---: | :---: | :---: | :---: | :---: |
| a. $\rightarrow$ əŋ |  |  |  | * | * |
| b. on |  | 上 | *! |  |  |
| c. y |  |  |  |  |  |
| d. ${ }^{\text {g }}$ |  |  |  |  |  |

High vowels are not considered to be inserted in example (68) because we have known that both of $*[\mathrm{ug}]$ and $*[\mathrm{yn}]$ are ill-formed rhymes through the analysis in 4.2.1.2. Even though [in] can be surfaced out, it could be interpreted that [iy] will not be the optimal candidate because of its violation of AGREE[bk]-V $\mathrm{H}_{\mathrm{H}} \mathrm{N}$. However, if we consider the other pattern of the same phenomenon, i.e. $/ \mathrm{n} /$ surfaced as $[\mathrm{m}]$, inserting a high vowel could be a possible strategy to repair this ill form. To avoid this prediction, constraint (69) is developed, which is also high-ranked, as tableau (70) shows. The Hasse Diagram is illustrated in (71).
(69) Dep-HiV

Assign one violation mark for every high vowel in the output which does not have a correspondent input.
(70) $/ \mathrm{n} / \rightarrow[\mathrm{m}]$

| In/ | Nuc | Ident[syl] | Dep-HIV | Dep-ə |
| :--- | :---: | :---: | :---: | :---: |
| a. $\rightarrow$ ən |  |  |  | $*$ |
| b. n | $*!$ |  |  |  |
| c. n |  | $*!$ |  |  |
| d. in |  |  | $*!$ |  |

(71) The Hasse Diagram of the inserted [ $\rho$ ] as a nucleus


As (71) shows, Dep-a is dominated by Nuc, Ident[syl] and Dep-HIV because only Dep-a is violable when it comes to a rhyme composed of only a nasal consonant in the input.

### 4.2.3 Low Vowel Preceding a Nasal Coda

### 4.2.3.1 Low Vowel Backing

In 4.2.1.1, it has been stated that high vowels should agree with the nasal coda in [back] feature while mid vowels do not have this restriction. This is proved by the fact that the ill form *[un] is repaired by inserting an [ə] in between and lead to the output
[wən], where the vowel and the nasal coda still do not agree in [back] with each other. Then what about low vowels? To answer this question, we just have to take a look at the patterns [an] and [ay]. Obviously, the low vowel is surfaced as [-back] when preceding $[\mathrm{n}]$ and $[+\mathrm{back}]$ when preceding $[\mathrm{y}]$. By this fact, it is known that low vowels, like high vowels, should also agree with the nasal coda in [back], but the strategy to repair the ill forms is different. As the strategy for low vowels to agree with the glide coda, they also change the value for the [back] feature to agree with the nasal coda.

In fact, it can be generalized that vowels except for mid ones should agree with the nasal coda in [back] feature, and the best strategy to repair the ill forms is to insert an [ə], unless the vowel cannot be glided (i.e. low vowels) and thus causes a complex nucleus. The analysis can be supported by the constraint ranking as below shows, where the lower index $L$ stands for low:

Agree[bk]-V $/$ /LN, *Comp-Nuc $\gg$ Ident[F]-V $\gg$ Dep-ə
However, it is not natural for a constraint to restrict the nasal coda to agree with its preceding "high or low" vowels with the mid vowels excluded. Therefore, here we tend to define two different constraints to regulate a high vowel and a low vowel to agree with the nasal coda in [back] feature respectively. One is Agree[bk]- $\mathrm{V}_{\mathrm{H}} \mathrm{N}$, which has been developed previously, and the other is illustrated in (73). The descriptive generalization is developed in (72), and an example is displayed in tableau (75). Constraint (74) is to ensure that the low vowel is not raised in this environment. As what is mentioned above, *Comp-Nuc is also active here for prohibiting an inserted [ə] causing a complex nucleus. Ident[bk]-N is also listed in the tableau to ensure that the nasal coda does not change.
(72) The descriptive generalization for low vowel backing

A low vowel should agree with its nasal coda in [back] feature. This requirement is enforced by vowel backing.
(73) Agree[bk]-VLN

Assign one violation mark for every syllable whose low vowel and nasal coda do not agree in the [back] feature.
(74) Ident[low]-V

Assign one violation mark for every vowel which has a different value for the [low] feature from its input.
(75) $/ \mathrm{ay} / \rightarrow[\mathrm{ay}]$

| /ay $/$ | Agree <br> $[\mathrm{bk}]-V_{\mathrm{L}} \mathrm{N}$ | *Comp-Nuc | Ident[bk]-N | Ident[low]-V | Ident[F]-V |
| :--- | :---: | :---: | :---: | :---: | :---: |
| a. $\rightarrow$ ay |  |  |  |  | $*$ |
| b. an | $*!$ | Che | hi |  |  |
| c. an |  |  | $*!$ |  |  |
| d. aəy |  | $*!$ |  |  |  |
| e. en |  |  |  | $*!$ |  |

When analyzing the phenomenon of inserting an [ə] between a high vowel and a nasal coda in 4.2.1.1, we ranked Ident[bk]-N and Ident[F]-V the same high because we had not had any evidence to compare these two constraints yet. In this part, it is observed that changing the feature of the vowel causes a better output than changing
the feature of the nasal coda. Therefore, that $\operatorname{Ident}[\mathrm{bk}]-\mathrm{N} \gg \operatorname{Ident}[\mathrm{F}]-\mathrm{V}$ is proved here.
If we take the patterns with a prenuclear glide into consideration, it is found that the low vowel still agrees with the nasal coda in [back] feature even though the prenuclear glide has a different value for the [back] feature from it, as tableau (76) shows. This proves that $A_{\text {gree }}[\mathrm{bk}]-\mathrm{V}_{\mathrm{L}} \mathrm{N} \gg$ Agree $\left.^{\mathrm{b} k}\right]-\mathrm{GV}^{10}$, while we have no evidence to prove which one of $A_{\text {gree }}[b k]-V_{H} N$ and $A_{\text {gree }}[F]-G V$ is ranked higher in 4.2.1.1. The Hasse Diagram is illustrated in (77) below.
(76) $/ \mathrm{jay} / \rightarrow[\mathrm{jay}]$

| /jay/ | Agree[bk]-V $\mathrm{V}_{\mathrm{L}} \mathrm{N}$ | Ident[low]-V | Agree[bk]-GV | Ident[F]-V |
| :---: | :---: | :---: | :---: | :---: |
| a. $\rightarrow$ jay |  |  | * | * |
| b. jay | (1) * |  |  |  |
| c. jen |  | *! |  |  |

(77) The Hasse Diagram of low vowel backing


Hasse Diagram (77) shows that when a low vowel precedes a nasal coda, it should agree with the coda in [back] feature even though it violates Agree[bk]-GV and

[^7]Ident[F]-V, as the example in (79) illustrates. Though both Agree[bk]-GV and Ident[F]V are violable here, that $\mathrm{A}_{\text {gree }}[\mathrm{bk}]-\mathrm{GV} \gg \operatorname{Ident}[\mathrm{F}]-\mathrm{V}$ has been proved in 4.1.1. Other constraints such as Ident[low]-V, *Comp-Nuc and Ident[bk]-N are inviolable, as example (75) shows.

### 4.2.3.2 Low Vowel Raising

Based on Diagram (77), we can successfully predict the following five patterns including [an], [ay], [jay], [wan] and [way], while *[jan], *[чan] and *[чay] are also predicted but fail to surface out. To figure out the problem, let's review the patterns of GVN where $\underline{\mathbf{V}}=\varnothing$ or mid vowel mentioned in 4.2.1.2, as is illustrated in (78).
(78) The combinations of $\underline{\mathbf{G V N}}$, where $\underline{\mathbf{V}}=\varnothing, \mathrm{e}, \partial, \mathrm{o}$.

| a. *n | *en $\boldsymbol{\text { n }}$ | * | e. *un | *wen | wən | *won |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| b. *y | *ey ə刀 |  | f. *uy | *wen | * wəy | *woy |
| c. in | ien *jən | *jon | g. yn | yen | чәп | * $¢ 0$ п |
| d. in | jen *jən | ${ }^{\mathrm{jon}}$ | h. *yy | *чеу | * чə | *чоу |

Theoretically, there should be accurately one allowed pattern in each item from (78a) to (78h), so there should be eight allowed rhymes in total while there are actually ten. Because the derivations of /uy/ to [oy] and /yn/ to [joy] have been discussed in 4.2.1.2 and $[\mathrm{in}]$ in ( 78 c ) and $[\mathrm{yn}]$ in ( 78 g ) are quite well-formed, the underlined patterns [jen] and [yen] are assumed to be the extra ones where the [e]'s are derived from low vowels. It is reasonable because *[jan] and *[чan] are both absent. The surface forms of the rhymes of $\underline{\mathbf{G V N}}$ structure, where $\underline{\mathbf{V}}$ is underlyingly a low vowel, are listed in (79)
below. This low vowel raising phenomenon has also been discussed by Hsieh (2012) (see 2.2.2).
(79) The surface forms of the rhymes of $\underline{\mathbf{G V N}}$ structure, where $\underline{\mathbf{V}}$ is underlyingly /a/
a. $/ \mathrm{jan} / \rightarrow[\mathrm{jen}]$
b. $/ \mathrm{jay} / \rightarrow[\mathrm{jay}]$
c. /wan/ $\rightarrow$ [wan]
d. /way/ $\rightarrow$ [way]
e. $/$ цаn/ $\rightarrow$ [чen]
f. $/$ чау/ $\rightarrow \emptyset$

Via (79), it is found that the low yowel is only raised in (a) and (e), where the prenuclear glides are both [-back] and the codas are both [n]. Because the failure for $/ \longleftarrow a \mathrm{y} /$ to surface is asymmetric, the null in ( f ) is considered to be an accidental gap.

As Hsieh (2012) assumes, the phenomenon of low vowel raising is considered a process of assimilation. However, his analysis where [a] is viewed as [+back] contradicts with the analysis of this thesis. Therefore, we assume that the low vowel is raised because of the [low] feature. Since this process is only triggered when there are a prenuclear glide and a nasal coda, we develop a constraint as (80) below. At the same time, this process, of course, violates a faithfulness constraint Ident[low]-V.
(80) Agree[low]-GVN

Assign one violation mark for every syllable whose prenuclear glide, vowel and nasal coda do not agree in the [low] feature.

However, there are no other [+low] segments in Taiwanese Mandarin except for the low vowels [a] and [a]. This indicates that all of the inputs in list (79) should undergo a low vowel raising process, which is not the fact. Therefore, it is necessary to explain why /jay/, /wan/ and /way/ do not undergo a low vowel raising process.

To observe (79a) and (79e), where $/ \mathrm{a} /$ is raised as [e], it is found that they have one thing in common. In these two patterns, the three segments of each rhyme are all [back]. By this fact, we can develop a preliminary descriptive generalization, as is shown in (81). Constraint (82) is developed based on (81).
(81) A preliminary descriptive generalization for low vowel raising

The segments in a GVN structure should agree in the [low] feature, unless they do not agree in the [back] feature. This requirement is enforced by vowel raising.

## (82) Agree[bk]-GVN

Assign one violation mark for every syllable whose prenuclear glide, vowel and nasal coda do not agree in the [back] feature.

According to (81), it is known that if a rhyme violates Agree[bk]-GVN, then it should not undergo the low vowel raising process; in other words, it should satisfy Ident[low]-V. Agree[bk]-GVN and Ident[low]-V should not be violated at the same time. Therefore, a local conjunction is developed, as in (83). Tableaux (84-85) illustrate how these constraints work.
(83) $\left\{\operatorname{Ident}[\text { low]-V \& Agree[bk]-GVN }\}_{\sigma}\right.$

Assign one violation mark for every syllable whose vowel has a different value for the [low] feature from its input and does not agree with either its prenuclear glide or nasal coda in [back] feature at the same time.
(84) $/ \mathrm{jan} / \rightarrow[\mathrm{jen}]$

| $/$ jan $/$ |  <br> Agree[bk]-GVN $\}_{\sigma}$ | Agree[low]-GVN | Ident[low]-V |
| :--- | :---: | :---: | :---: |
| a. $\rightarrow$ jen |  |  | $*$ |
| b. jan |  |  | $*$ |

(85) /wan/ $\rightarrow$ [wan]

| /wan/ |  <br> Agree[bk]-GVN\}。 | Agree[low]-GVN | = Ident[low]-V |
| :---: | :---: | :---: | :---: |
| a. $\rightarrow$ wan | ${ }^{2}$ | * |  |
| b. wen | *! |  | * |

Both of the examples in tableaux $(84-85)$ contain an alveolar coda [ n ]. When it comes to the rhymes with the velar coda [ y$]$, the constraints involved low vowel backing, which are discussed in 4.2.3.1, should also be considered, for the low vowel should be retracted as [a] in the environment. In list (79), it is observed that the low vowels in /jan/ and /чan/ are raised, while those in/jay/ and/way/ are retracted, and the low vowel in /wan/ does not undergo any process. None of them is raised and retracted at the same time. Based on this fact, a revised descriptive generalization isdeveloped in (86) below, and the local conjunction (87) is developed according to (86).
(86) The revised descriptive generalization for low vowel raising

A low vowel is raised in a GVN structure where the segments agree in the [back] feature, unless the low vowel should be retracted to agree with the nasal coda in [back].
(87) $\{\operatorname{Ident}[l o w]-\mathrm{V} \& \operatorname{Ident[F]-V}\}_{\sigma}$

Assign one violation mark for every syllable whose low vowel has different value for its [back] or [round] feature from its input, and simultaneously has a different value for its [low] feature from its input.

Based on the description in (86), it is known that a low vowel cannot be raised and retracted simultaneously. When a low vowel is in the environment where it should be both raised and retracted, it chooses to be retracted to satisfy Agree $^{\operatorname{bbk}]-\mathrm{V}_{\mathrm{L}} \mathrm{N} \text { instead of }}$ to be raised to satisfy Agree[low]-GVN. Therefore, Agree $^{\text {Gkj}]-V_{L} N}$ should be ranked higher than Agree[low]-GVN. The local conjunctions in (83) and (87) are never violated, so they are ranked the highest along with Agree $[b k]-V_{\mathrm{L}} \mathrm{N}$. Tableaux (88) and (89) illustrate how this constraint ranking works, and the Hasse Diagram is illustrated in (90).
(88) $/ \mathrm{jan} / \rightarrow[\mathrm{jay}]$

| /jan/ | $\{\operatorname{Ident}[10 w]-\mathrm{V} \text { \& _ }\}_{\text {c }}$ |  | Agree <br> $[\mathrm{bk}]-\mathrm{V}_{\mathrm{L}} \mathrm{N}$ | Agree <br> [low]-GVN | Ident <br> [low]-V |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | Agree[bk]- <br> GVN | Ident[F]-V |  |  |  |
| a. $\rightarrow$ jay |  |  |  | * |  |
| b. jay |  |  | *! | * |  |
| c. jen | *! |  |  |  | * |

(89) /way/ $\rightarrow$ [way]

| /way/ | \{IDent[low]-V \&_\} AGREE[bk]- GVN IDEnt[F]-V GVN | Agree [low]-GVN | Ident <br> [low]-V |
| :---: | :---: | :---: | :---: |
| a. $\rightarrow$ way | $\square$ | * |  |
| b. way | - | * |  |
| c. wey | *! |  | * |
| d. wə | *! |  | * |

(90) The Hasse Diagram of low vowel raising

$$
\left.\left\{\operatorname{Ident}[l o w]-\mathrm{V} \& \text { Agree }^{2} \mathrm{bk}\right]-\mathrm{GVN}\right\}_{\sigma} \quad\left\{\operatorname{Ident}[\text { low]-V \& Ident[F]-V }\}_{\sigma}\right.
$$



Though not illustrated in tableaux (88) and (89), that the local conjunctions dominate their conjoining members is universal. There are still some domination relations having been proved in 4.2.3.1. First, Both Agree[bk]-V ${ }_{\text {L }}$ and Ident[low]-V dominate $A_{\text {gree }}[b k]-G V$. Second, Agree[bk]-GV dominates Ident[F]-V. Since Agree[bk]-GVN has a stricter definition than Agree[bk]-GV, the former also dominates the latter.

### 4.2.4 Summary

From section 4.2.1 to 4.2.3, we have discussed all the patterns with a nasal coda by different underlying structures. First we discussed the patterns that are underlyingly high vowels preceding the nasal, and then the analysis goes on to the patterns which has only a nasal segment in the input. Last, we discussed the low vowels preceding a nasal coda. The results are concluded as below:
(i) An [ə] is inserted when a high vowel and its following nasal coda do not agree in the [back] feature. (see 4.2.1.1)
(ii) The rule in (i) is banned when the process leads to [əŋ]. /iy/ surfaces faithfully as [in] to avoid [əŋ] while $/ \mathrm{yy} / \mathrm{splits}[\mathrm{y}$ ] to have [joy] in the output. (see 4.2.1.2)
(iii) Two high back segments cannot be adjacent in the rhyme domain. /uy/ is surfaced as [oy] to avoid this. (see 4.2.1.2)
(iv) An [ə] is inserted when there is no vowel in the rhyme domain. (see 4.2.2)
(v) A low vowel should agree with its nasal coda in [back] feature. (see 4.2.3.1)
(vi) A low vowel is raised only when the preceding glide and the following nasal are both [-back], for the segments in the rhyme should agree in [back] after vowel raising and a low vowel cannot be raised and retracted at the same time. (see 4.2.3.2)

Through these analyses, we have developed a Hasse Diagram for each phenomenon. Like what has been done in 4.1.4, an overall Hasse Diagram is to be developed. In the discussion of the rhymes with a nasal coda, much more constraints are active in not only one analysis, so the Hasse Diagram is more complicated, as shown in (91).
(91) The Hasse Diagram of rhymes with a nasal coda


Diagram (91) successfully predicts all the rhymes with a nasal coda in Taiwanese Mandarin except for the absence of $*[ч \boxed{\square}]$. As what is assumed when discussing the issue of low vowel raising, *[ $[\square \square]$ is considered an accidental gap because the absence is asymmetric.

## Chapter 5

## Conclusion

Through the analysis in Chapter Four, it is concluded that the present and absent patterns of Taiwanese Mandarin rhymes are mostly regular, and several constraints are active to restrict the phonotactics. As what has been mentioned, some constraints are not only active in a single phenomenon. In section 4.1.4, we have illustrated a Hasse Diagram to show all the active constraints restricting the patterns without a nasal coda. On the other hand, in section 4.2.4, a Hasse Diagram is displayed to illustrate the constraints about the patterns with a nasal coda. Though these two Hasse Diagrams have respectively shown the constraint rankings for patterns with and without a nasal coda, some constraints are active in both kinds of patterns. Therefore, these two Hasse Diagrams can be combined into one to illustrate the constraint ranking for Taiwanese Mandarin rhymes. After all, theses constraints are restricting the phonotactics of a certain dialect of a language, and an overall Hasse Diagram helps clarify the domination relations between every pair of constraints. For example, there is no evidence to prove that Max $\gg \operatorname{Ident}[\mathrm{F}]-\mathrm{V}$ when discussing the patterns without a nasal coda while it is proved when it comes to the patterns with a nasal coda for the fact that MAX >> Dep-e, Dep-e >> Agree[bk]-GV and Agree[bk]-GV >> Ident[F]-V. The overall Hasse Diagram is illustrated in (1).
(1) The Hasse Diagram of Taiwanese Mandarin rhymes


Through the analyses and the constraint ranking, all the inexistent rhymes can be ruled out except for $*[ч а]$ and $*[ч а \square]$, which is respectively mentioned in 4.1.4 and 4.2.4. Due to their asymmetric absence, these two inexistent patterns are considered accidental gaps. Previous research has regarded much more absent rhymes as accidental gaps, including *[ча], *[чаך], *[чəŋ] (Lin, 2007), *[wi], *[ju], *[чən] and *[jən] 78
(Duanmu, 2007). However, through our analyses, it is found that the absence of most of these patterns are in fact systematic and can be explained. Only *[чa] and *[чaŋ] are considered to be accidental gaps.

This thesis discusses the phonotactic restrictions of Taiwanese Mandarin rhymes and finds several constraints to rule out the absent patterns. However, some allowed rhymes can follow certain onsets but cannot follow others. For example, we have proved that [jay], as the optimal output of /jay/, is existent as a rhyme in Taiwanese Mandarin. However, although it can follow onsets like [1] and [ $\mathrm{t}^{\mathrm{h}}$ ] to form liang [ljay]
 follows the allowed onsets [ p ] and $\left[\mathrm{t}^{\mathrm{h}}\right.$ ], are absent. The phonotactic restrictions of Taiwanese Mandarin "syllables" is therefore a further issue to be researched based on this thesis.


## REFERENCES

Chung, Raung-Fu (1995). The Segmental Phonology of Southern Min in Taiwan. Taipei: Crane Pub. Co.

Duanmu, San. (1990). A Formal Study of Syllable, Tone, Stress, and Domain in Chinese Languages. Ph.D. dissertation, MIT.

Duanmu, San. (2007). The Phonology of Standard Chinese (Second Edition). New York: Oxford University Press.

Flemming, Edward. (2003). The Relationship between Coronal Place and Vowel Backness. Phonology 20(3), 335-373.

Fon, Janice, Jui-mei Hung, Yi-Hsuan Huang and Hui-ju Hsu. (2001). Dialectal Variations on Syllable-Final Nasal Mergers in Taiwan Mandarin. Language and Linguistics, 12(2), 273-311.

Hsiao, Yuchau E. (2000). Optimal Tone Sandhi in Taiwanese. Chinese Studies 18:25-40. [In Chinese]

Hsieh, Feng-fan. (2012). Low Vowel Raising in Sinitic Languages: Assimilation, Reduction, or Both? Language and Linguistics 13(4), 583-623.

Hsu, Dong-Bo. (2009). The Syllable in Old Chinese: Sub-Syllabic Processes, Syllable Structure, and the Status of Medial Glides. J East Asian Linguist 18:361-395

Ito, Junko and Armin Mester. (1998). Markedness and Word Structure: OCP Effects in Japanese. ROA-255-0498.

Kager, Rene. (2004). Optimality Theory. New York: Cambridge University Press.
Leben, Will. (1973). Suprasegmental Phonology. Doctoral dissertation, MIT.
Lee, Chun-Chih. (2001). Medials and the Evolution of the Initial System in Chinese. Chinese Phonology 11:69-108

Lin, Yen-Hwei. (2007). The Sounds of Chinese. New York: Cambridge University Press. Lin, Yen-Hwei. (2015). OT Factorial Typology and the Typology of Mandarin Mid Vowel Assimilation. Language and Linguistics 16(6), 769-789.

McCarthy, John J and Alan Prince. (1995). Faithfulness and Reduplicative Identity. University of Massachusetts Occasional Papers in Linguistics 18, 249-384.

McCarthy, John J and Alan Prince. (1999). Faithfulness and Identity in Prosodic Morphology. Cambridge: Cambridge University Press.

McCarthy, John J. (2008). Doing Optimality Theory: Applying Theory to Data. Malden, MA and Oxford: Blackwell Publishing.

Prince, Alan., and Paul Smolensky. (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Malden, MA: Blackwell.

Yip, Virginia and Stephen Matthews. (2006). Cantonese: A Comprehensive Grammar (Routledge Comprehensive Grammars). London: Routledge.


[^0]:    ${ }^{1}$ Whether the prenuclear glides (i.e. the $\underline{\mathbf{G}}$ of $\underline{\mathbf{G V X}}$ ) belong to the onset or the rhyme has been quite controversial. Duanmu (1990) proposes that a prenuclear glide is actually the secondary feature of the onset. For example, kuan 'close' is actually pronounced [ $\left.\mathrm{k}^{\mathrm{W}} \mathrm{an}\right]$, where $[\mathrm{k}]$ has a secondary feature of [+round]. However, Hsu (2009), on the other hand, thinks that [j] belongs to the rhyme while [w] belongs to the onset. Lee (2001) also states that prenuclear glides of Mandarin possess a very special position, whose phonological feature is contradicted to Chinese phonological system. Here I clarify that the domain of rhyme in this thesis includes prenuclear glides, which will later be elaborated in 1.2.1.
    ${ }^{2}$ Syllables like [e] and [o] appear as interjections or particles in Mandarin. However, they never appear in other lexical words. The clarification is elaborated in section 3.2.1.
    ${ }^{3}$ The inventories of each position will be displayed later in section 3.1.

[^1]:    ${ }^{4}$ To analyze more accurately, the off-glides [j] and [w] are treated as codas in this thesis instead of parts of the nuclei. However, the term "diphthong" is still used to call a sequence of a vowel and an off-glide. The use of the term "triphthong" is in the same way. For example, [ej] and [aw] are diphthongs; [wej] and [jaw] are triphthongs.

[^2]:    5 Asp stands for aspiration. In OT, if there are two winners among the candidates, they might indicate different speech variations

[^3]:    ${ }^{6}$ Lin defines the term "rime" as the structure of VX. Therefore, her Rime-Harmony constraint has the same definition as Duanmu's NC-Harmony in (9) and (10).

[^4]:    ${ }^{7}$ Based on Hashimoto's theory (see 3.1), it is the prenuclear glide instead of the postnuclear one that spreads to the $\underline{\mathbf{V}}$ position. Therefore, the sequence of two high vocoids is more properly transcribed as "vowel + glide." However, here we transcribe a sequence of two high vocoids as "glide + vowel", for it is more traditional and easier to read. Also, the form of "glide + vowel" is assumed to be the underlying form of a sequence of two high vocoids. To avoid being controversial both of the transcriptions are available in chart $(2-5)$.

[^5]:    ${ }^{8}$ The F in the bracket stands for Feature, which can be substituted into [back] or [round].

[^6]:    ${ }^{9}$ All of the examples in (31) have a lax [r] following [w]. However, vowels do not contrast in tenseness in Mandarin, so the output will not be influenced by the tenseness of the vowel in the input.

[^7]:    ${ }^{10}$ Theoretically, whether Agree $[\mathrm{bk}]-\mathrm{V}_{\mathrm{L}} \mathrm{N}$ dominates Agree $\left.^{2} \mathrm{rd}\right]-\mathrm{GV}$ has no evidence, for candidates with a low rounded vowel are ruled out by a high-ranked constraint Рнолотастісs. Therefore, we only confirm that Agree $[\mathrm{bk}]-\mathrm{V}_{\mathrm{L}} \mathrm{N}$ dominates Agree[bk]-GV.

