PHONOLOGICAL ANALYSIS OF HONGKUI To

by

Xu Zhuo

A dissertation submitted to Guangdong University of Foreign Studies

in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy

Guangdong University of Foreign Studies

April 2001

To my parents

who nurtured me

and

My husband

who has supported me with unconditioned love

Abstract

This dissertation discusses the phonological system of Hongkui To (henceforth HT) within the framework of Government Phonology (henceforth GP). The discussion is divided into the following aspects: the phonological template, the nuclear system, the onset system, inter-constituent interaction and the tone system.

First, in chapter one, I make a brief introduction on the basic notions of GP, which includes the element theory, theory of constituent structure, theory of transconstituent structure, government, Empty Category Principle and tone theory.

Chapters two to six are considered the core part of the whole dissertation. In Chapter Two, I discuss the phonological template of HT, proving that it is a templatic language. The basic template of HT is two sets of onsets and nuclei, i.e. $O_1N_1O_2N_2$, on the basis of which an extended template, $O_1N_1O_aN_aO_2N_2$, can be derived. P-licensing Constraint and other related constraints will also be discussed.

Chapter Three focuses on the nuclear system of HT. The concept of licensing constraints in general is introduced first. Then the licensing constraints of the nuclear system at the lexical level of HT are established, from which all and only

the seven lexical nuclei of HT can be derived. The phonological representations for the lexical nuclei are provided accordingly.

The phonological representations of onsets and their behavior are discussed in the fourth chapter. I propose the lexical representation of HT onsets in terms of elements. Specific constraints on onsets and their behavior are also included.

To provide a comparatively complete picture of the phonological system of HT, Chapter Five provides discussion specifically on the interaction between constituents. There are three sections, working on the interaction between nuclei, the interaction between onsets and the interaction between onset and nucleus respectively. Some of the constraints are found echoing with each other, which provide evidence to the discussion and claims.

The discussion of the tone system is provided in Chapter Six. HT is found to be a language of complex tone system. Constraints are proposed to derive the five tonal patterns of the language. It is discovered that the spreading always goes from right to left, there can be at most one tonal element in a tonal patterns, and that the final position of the tonal patterns is the head.

Table of Contents

Chapter One

INTRODUCTION

Chapter Two

THE PHONOLOGICAL TEMPLATE

Chapter Three

THE NUCLEAR SYSTEM

Chapter Four

THE SYSTEM OF ONSETS AND THEIR REPRESENTATIONS

Chapter Five

INTER-CONSTITUENT INTERACTION

Chapter Six

TONE SYSTEM

Abbreviations and Transcriptions

1. Abbreviations

BJ	Beijing
Co	Coda
ECP	Empty Category Principle
GP	Government Phonology
HT	Hongkui To
KLV	Kaye, Lowenstamm & Vergnaud
N	Nucleus
O	Onset
OCP	Obligatory Contour Principle
P_0	P-zero
P^N	Nuclear projection
R	Rhyme
UG	Universal Grammar
LC	Licensing Constraint

2. Transcription

The alphabetic system of transcription followed in this dissertation is a Zhuang font created specially for the phonological research of Zhuang language. However, the reader should not find it difficult to understand as it follows the general convention in phonetic transcription. Please note that the use of phonetic transcription in connection with phonological representations is merely a convenience to assist those not familiar with the theory in Government Phonology.

Acknowledgements

I never thought that I would be a PhD, from my childhood to even around five years ago. Yet today, an unexpected dream has come true. With years of not so common experiences in mind, I am very pleased and honored to take this opportunity to thank people who have supported and assisted me.

My supervisor, Professor Jonathan Kaye, deserves special thanks for his enduring guidance, continuous encouragement, and great care. Needless to say, the thesis would not have been what it is now, nor would it have been completed. It has been a great privilege to be his student and to work with him. I have an extreme admiration for his rigorous effort and great knowledge of phonology. In the process of writing and correcting this thesis, he provided not only penetrating criticism and comments but also admirable enlightment.

I also would like to express my special thanks to Professor Gui Shichun, who offered me the opportunity to be a PhD student by accepting my application at the first place. Although I chose phonology instead of psycholinguistics later, he has always shown his concern and support.

Professor Ning Chunyan also deserves my special thanks for his part. I have immensely benefited from the discussions with him, which in fact provided me the

basis of theoretical linguistics and led to my later choice of phonology.

I shall never forget the generosity, enthusiasm and friendship of a number of fellow research students in the PhD Program, Guangdong University of Foreign Studies. They made my years in Guangzhou, a city more than 3,000 km from my hometown, enjoyable, fruitful, and intellectually stimulating. I feel grateful to Zhang Wenzhong for his always being there for me, He Jianing, for his humor and staying up for two nights with me, doing the statistics assignment in the computer room, Yang Shouxun, for his invaluable trust and support in the most difficult time, and a number of MA students, who have been such good friends and listeners. The girl talks will always be a nice memory in my life.

I would also like to express thanks to Huang Meiying, who provided most of the raw data in this research, and her family in Hongkui Village, who provided assistance in my fieldwork.

My parents, who although don not really understand what I have been working on, have always offered me financial and spiritual support with faith and love.

Finally, I would like to thank my husband for his great support and faith, enduring patience and indulgence over the course of this long thesis. I would like to share this achievement with him.

宏魁土语音系分析

许卓（Xu Zhuo）

本篇论文对宏魁土语进行了一个较为全面的音系分析研究，研究所采用的理论框架为管辖音系理论。论文共分为以下几个方面：音系模板、韵核系统、首音系统、成分间作用和音调系统。

第一章首先简单介绍了有关管辖音系理论的一些基本概念，包括元素理论、成分结构理论、跨成分结构理论、管辖、空范畴原则和音调理论。

第二章到第二章应该是整篇论文的核心部分。第二章探讨的是宏魁土语的音系模板，证明了这是一种模板语言。宏魁土语的基本模板是两对首音和韵核，也就是 $O_1N_1O_2N_2$；同时，在这一模板基础上还发现了一个扩充的模板——$O_1N_1O_aN_aO_2N_2$。这一章还讨论了 P-允准制约和其它相关的制约。

第三章集中讨论的是宏魁土语的韵核系统。笔者首先概括地介绍了允准制约的概念，然后建立了宏魁土语词汇层面上的韵核系统的允准制约。通过这些允准制约可以推导出，也只能推导出宏魁土语的七个词汇韵核。在此基础上，还提供了这些词汇韵核的音系表达式。

第四章讨论的是首音的音系表达式和它们的表现，并用元素理论提出了宏魁土语首音的词汇表达式。这一章还包括了首音的具体的制约和相应表现。

为了对宏魁土语的音系系统提供一个相对完整的的描述，笔者在第五章中特别讨论了成分之间的相互作用。这一章主要分为三个部分，分别讨论韵核之间的相互作用，首音之间的相互作用和首音与韵核之间的相互作用。其中的有些制约彼此响应，为其中的讨论和观点提供了证据。

第六章讨论的是音调系统。论文研究发现宏魁土语是一个复杂音调系统的语言，并提出了可以得出此种语言五个音调模式的制约条件。另外，这一章中还有如

下发现：在宏魁土语的音调系统中，延伸的方向总是从右向左；在一个音调模式中最多只能有一个音调元素；音调模式中的最后一个位置是首部。

关键词：管辖音系理论，宏魁土语，音系模板，扩充模板，延伸，音调理论，普通话，广东话。

Abstract

Xu Zhuo

This dissertation discusses the phonological system of Hongkui To (henceforth HT) within the framework of Government Phonology (henceforth GP). The discussion is divided into the following aspects: the phonological template, the nuclear system, the onset system, inter-constituent interaction and the tone system.

First, in chapter one, I make a brief introduction on the basic notions of GP, which includes the element theory, theory of constituent structure, theory of transconstituent structure, government, Empty Category Principle and tone theory.

Chapters two to six are considered the core part of the whole dissertation. In Chapter Two, I discuss the phonological template of HT, proving that it is a templatic language. The basic template of HT is two sets of onsets and nuclei, i.e. $O_1N_1O_2N_2$, on the basis of which an extended template, $O_1N_1O_aN_aO_2N_2$, can be derived. P-licensing Constraint and other related constraints will also be discussed.

Chapter Three focuses on the nuclear system of HT. The concept of licensing constraints in general is introduced first. Then the licensing constraints of the nuclear system at the lexical level of HT are established, from which all and only

the seven lexical nuclei of HT can be derived. The phonological representations for the lexical nuclei are provided accordingly.

The phonological representations of onsets and their behavior are discussed in the fourth chapter. I propose the lexical representation of HT onsets in terms of elements. Specific constraints on onsets and their behavior are also included.

To provide a comparatively complete picture of the phonological system of HT, Chapter Five provides discussion specifically on the interaction between constituents. There are three sections, working on the interaction between nuclei, the interaction between onsets and the interaction between onset and nucleus respectively. Some of the constraints are found echoing with each other, which provide evidence to the discussion and claims.

The discussion of the tone system is provided in Chapter Six. HT is found to be a language of complex tone system. Constraints are proposed to derive the five tonal patterns of the language. It is discovered that the spreading always goes from right to left, there can be at most one tonal element in a tonal patterns, and that the final position of the tonal patterns is the head.

Keywords: Government Phonology (GP), Hongkui To (HT), phonological template, augmented template, spreading, tone theory, Mandarin, Cantonese

Chapter One

Introduction

1.1 Research Orientation

The proposed research aims at doing a systematic phonological analysis of the Zhuang language spoken in the village of Hongkui Tiandeng County, Guangxi Zhuang Autonomous Zone, China. This study will be the first to analyze the Zhuang language in the perspective of theoretical linguistics, even more so within the framework of Government Phonology (henceforth GP), which is based on principles and parameters rather than rules or features. Although Zhuang has been studied for decades, most of the research done so far is descriptive. I will concentrate on five major issues of this language: (i) its phonological template, (ii) the nuclear system, (iii) the onset system, (iv) the inter-constituent interaction, and (v) the tone system.

The aim of this dissertation is three-fold. First, as the focus of this research falls on the variety of Zhuang spoken in Tiandeng County, which is a variety of

Southern Zhuang. This phonological research of the specific variety of Zhuang in question is a pioneer study in the field, providing the first linguistic data on the language. Secondly, by applying GP as the theoretical framework, the research will hopefully make its contribution in examining and advancing the theory of GP. Finally, it is expected to draw more attention on the theoretical study of these "minority languages".

1.2 Rationale

As this dissertation works on one of the "minority languages" spoken in China, it is necessary to have a brief introduction and description of the Zhuang language and the state of art of related studies, which will help illustrate the *raison d'être* of the present research.

1.2.1 About Zhuang

Zhuang is a minority language spoken by the Zhuang people in China. It belongs to the Tai linguistic family. There are two main subgroups of the Tai linguistic family in China, a) Zhuang spoken in most areas of GuangXi Zhuang

Autonomous Zone, and b) BuYi or Man spoken in Southern GuiZhou. There are two dialects in Zhuang—Northern Zhuang and Southern Zhuang. Southern Zhuang is generally spoken in NanNing City and area of the West of the Right River in GuangXi, and Southern YunNan, while Northern Zhuang is generally spoken in Northern GuangXi. The varieties of the two dialects are called To by convention (Liang & Zhang, 1996). There are seven varieties of To in Northern Zhuang, namely, GuiBei, LiuJiang, HongShui River, GuiBian, Northern Yi, the Right River, in GuangXi Zhuang Autonomous Zone and QiuBei in YunNan Province; and five varieties in Southern Zhuang—Southern Yi, the Left River, DeJing, in GuangXi, and YanGuang and WenMa in YunNan Province. The focus of this research falls on the To spoken in Hongkui Village, TianDeng County, which is a variety of the Left River To of Southern Zhuang. This is why I call it Hongkui To (henceforth HT) instead of Hongkui Zhuang throughout the dissertation, as it is the convention of both the Zhuang people and linguistic studies.

There are certain differences between the two groups of Zhuang, in phonology, the lexicon and grammar, which cause quite some difficulties of mutual comprehension. The most prominent difference is in the phonology. In a word, Zhuang has rich linguistic data which may well help to discover some of the mysteries of human languages, and this very first formal theoretical study on HT is expected to draw the curtain for the comprehensive linguistic research on Zhuang in the future.

1.2.2 State of art

Minority languages in China have been widely studied. A huge amount of fieldwork has been carried out throughout the country, including the investigation of To. Of the specific variety of To mentioned above, i.e. the one spoken in HongKui Village of TianDeng County, however, no fieldwork has ever been done so far as we know. The closest varieties of the language that have been investigated are from two areas of the same county, JiangNan Village (Qin, 1996) and the Lichuan area, which differ from the variety in question. Thus, doing phonological research on the specific variety of To will be the pioneer study in the field. On the other hand, although To has been studied for decades, most of the research done so far is still descriptive, thus a systematic theoretical study of the phonological system of the language can fill this gap.

1.3 Theoretical Framework

It has been mentioned at the beginning of this chapter that this research will be carried out in the framework of GP. The theory was first proposed by Kaye,

Lowenstamm and Vergnaud (1985, 1990). Other fundamental papers on GP include Kaye (1990, 1995), Charette (1988, 1990, 1991), Harris (1990, 1994), Harris and Kaye (1990), Yoshida (1990), etc. Inspired by the study of syntax, GP aims to look for the 'Universal Grammar' in phonology. Thus the theory is based on universal principles and language-specific parameters instead of features or rules, with the aim of providing non-arbitrary accounts for various phonological phenomena. One underlying assumption of GP is that words are not constituted of unorganized sequences of segments. This section will provide an introduction to the basic and major notions in GP. The details of some other aspects of GP will be presented in the body of the dissertation when they are relevant to the discussion.

1.3.1 Element theory

The basic components in GP are a set of elements, E, which is defined as follows:

(1) E = {A, I, U, H, L, ?}

There is also an identity element, usually represented as "_" in phonological expressions. Each of these elements is an independent, (potentially) interpretable phonological expression. The interpretation of each element, according to Kaye (2000a), "depends on (i) what phonological constituent dominates it, and (ii)

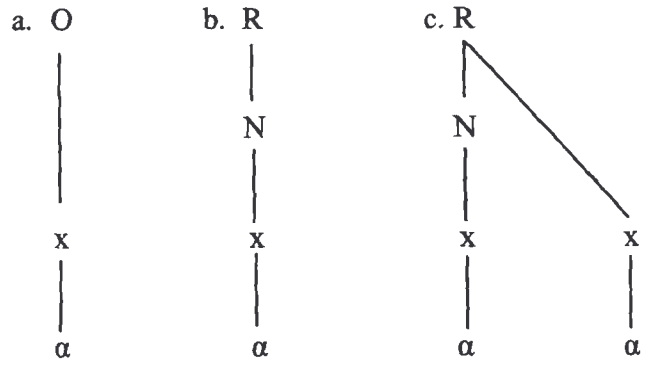
whether it occupies a head or operator position within a phonological expression". With these elements, the phonological expressions in all human languages can be represented, as all speech sounds are phonological expressions.

Each phonological expression is composed of two parts: a *head*, which may be null or has at most one element, and zero or more *operator*s. By convention, the operator(s) is/are usually put as the first member in a phonological expression, and the head, the second. Expressions headed by the identity element are called headless, while the others are called headed. There is licensing at various levels in GP. In a phonological expression, the head is said to license the operator(s). There are also licensing constraints, which are language-specific parameters regulating the combination of elements. Chapter Three will provide a detailed discussion on the notion of licensing constraints.

1.3.2 Theory of constituent structure

Above the level of elements is that of constituents. GP recognises three constituents—*onset*, *nucleus*, and *rime*, which contain elements associated to their skeletal points. The structures are shown overleaf.

(2)



(constituent level)

(skeletal level)

(segmental level)

In the above structures, the top tier is the constituent level. There are three types of "syllabic" constituent, i.e. onset (O), nucleus (N) and rhyme (R) respectively. The second tier is the skeletal points, which are represented by 'x's as a convention. The skeletal points link constituents and segments. The tier below the skeletal points is the segmental level for phonological expressions, which are composed of elements. It may be noted from the above structures that syllable and coda have no theoretical status in GP, though GP does recognize the post-nuclear rhymal position as in (2c), and a "syllable" can be interpreted as a sequence of onset-rhyme.[1] The next chapter on the phonological template of HT will provide detailed discussion on this.

As the name of the theory suggests, there is government between and within constituents. According to KLV (1990), governing relations are established at the level of lexical representation. Furthermore, a principle of grammar, the Projection Principle (KLV, 1990), ensures that there is no fundamental change in governing relations and their constituent structure from underlying representation

[1] See KLV (1990) for the arguments of rejecting them as constituents.

to the final output of a derivation. The phonology cannot manipulate governing relations. This is formulated in KLV (1990:221) as follows:

(3) Projection Principle

Governing relations are defined at the level of lexical representation and remain constant throughout a phonological derivation.

This principle precludes any changes in governing relations during the course of derivation, which means resyllabification is impossible. An onset may not change to a nucleus, or vice versa, as there are governing and licensing relationships between and within constituents.

For government within constituents, each constituent is considered as a governing domain, and the two principles that must be followed in constituent structure are *strict locality* and *strict directionality*.

(4) Conditions for Constituent Government

i. *Strict locality*: a governor and its governee must be strictly adjacent at the P_0 projection.

ii. *Strict directionality*: all constituents are head initial.

The above two principles define governing relations as binary and asymmetric, which means that one position within a governing domain governs, while the

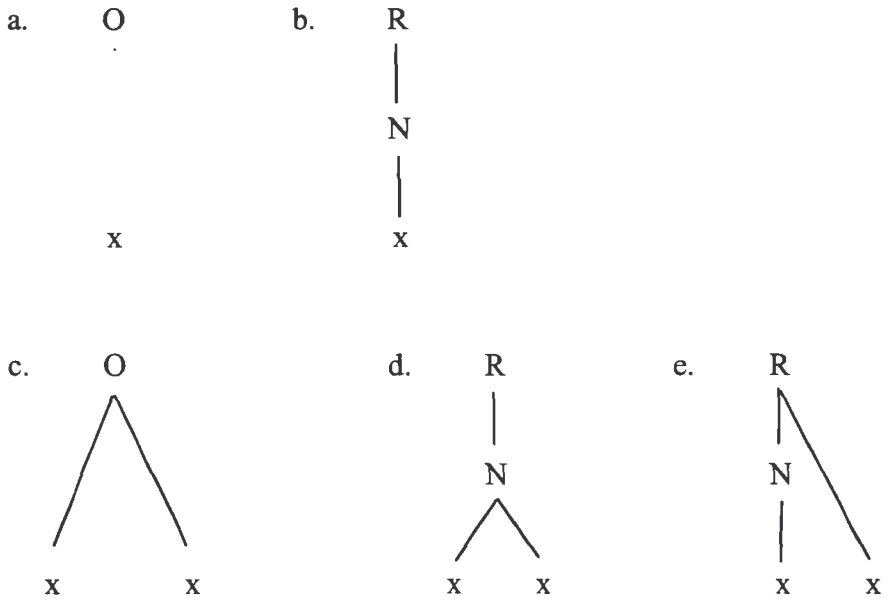
other is governed. The former is usually referred to as the governor, while the latter is known as the governee. From the above two principles, Binarity Theorem can thus be derived, which requires that all constituents are maximally binary, i.e., they may have only 0, 1 or 2 positions. Kaye (1990) thus proposed the following theorem.

(5) Binary Theory

All syllabic constituents are maximally binary.

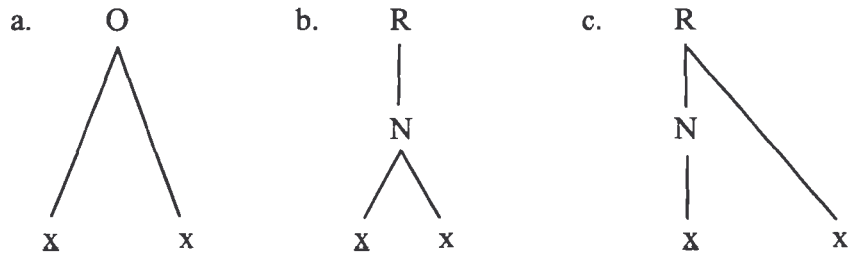
All the three constituents are supposed to be organized according to (4) and (5). They may or may not branch, depending on the specific constraints of the language. Therefore, there are five possible constituent structures allowed in GP shown overleaf.

(6) Constituent structures



Under constituent government, every branching constituent is a government domain. From the above principles, it is easy to derive that the left member of a branching constituent, as the head of each constituent, governs the right member, for the principle in (4) requires the constituent government to be universally head-initial. This is illustrated in the following structures.

(7) Constituent government



As the arrow indicates, in the above structures, the government goes from left to right, since the head of the domain, which is underlined, is on the left.

1.3.3 Theory of transconstituent structure

The level above constituent structure is that of the transconstituent structure, which occurs when the governor follows the governee, or when two skeletal points are dominated by different constituents. The direction of transconstituent government is just the opposite of that of the constituent government, which is also strictly directional, but head final in this case. Thus, there are also two conditions for transconstituent government as in constituent government:

(8) Conditions for tranconstituent government

i. *Strict locality*: a governor and its governee must be strictly adjacent.

ii. *Strict directionality*: the head is final.

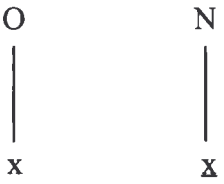
According to the principles in GP, every nucleus can and must license an onset to its left, and every rime must contain a nucleus as its head branch. The principle of onset licensing was proposed by Harris (1992:380) as overleaf.

(9) Onset Licensing

An onset head position must be licensed by a nuclear position.

This principle naturally leads to the conclusion that onsets must always be followed by a nucleus. This licensing can be illustrated as follows:

(10)



The above structure can also be an illustration of the governing relationship between an onset and a nucleus, as in transconstituent government, the head is final, i.e. on the right.

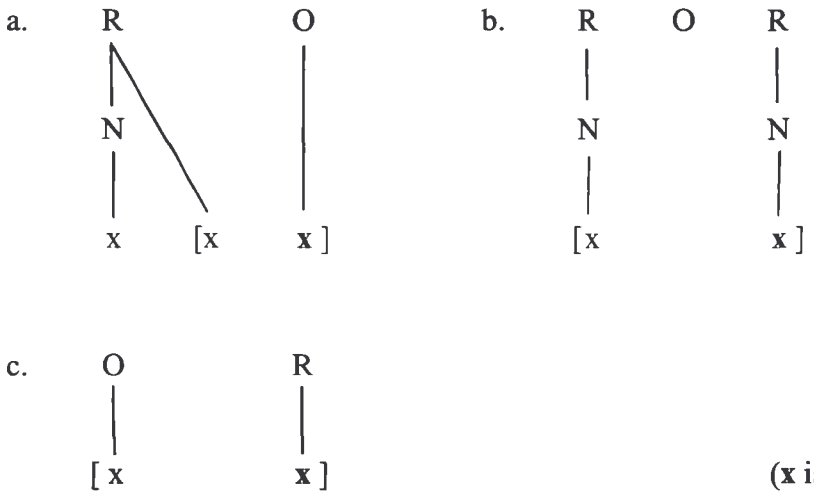
From the above theories of constituent structure and transconstituent structure, it is not difficult to see that a governing relation may hold between any two adjacent skeletal points, either constituent or transconstituent. Charette (1991) thus proposed two principles:

(11) i. Only the head of a constituent can govern.

ii. Only the nuclear head may govern a constituent head.

And three governing configurations were derived from the above two principles:

(12) Transconstituent governing domains



(**x** is the head.)

(Charette, 1991: 24)

1.3.4 Empty Category Principle

In addition to the above theories on elements, constituent and transconstituent structures, GP proposed a specific principle—the phonological Empty Category Principle (henceforth ECP), which is defined as follows (Kaye, 1995:295), in order to deal with the so-called epenthesis or syncope phenomenon in phonological derivation.

(13) Empty Category Principle

i. The Phonological ECP

A p-licensed (empty) category receives no phonetic interpretation.

ii. P-licensing

a. Domain-final (empty) categories are p-licensed (parameterized)

b. Properly governed (empty) nuclei are p-licensed

c. A nucleus within an inter-onset domain

(14) Proper Government:

α properly governs β if:

i. α and β are adjacent on the relevant projection,

ii. α is not itself licensed, and

iii. No governing domain separates α from β

In the ECP, proper government requires the absence of phonological material, i.e. elements, while p-licensing determines whether a part of the phonological string is audible or not. This implies that the licensing of word-final empty nuclei is parametric. On the one hand, in languages like English and Chinese, where domain-final nuclei may be p-licensed, words may appear to end in consonants. On the other hand, in other languages, like Japanese, which do not license word-final empty nuclei, p-licensing of a domain-final nuclear position is not allowed, thus these nuclei must be provided with phonetic content. Secondly, what

determines whether or not a word-internal empty nucleus may be left without phonetic manifestation is whether or not it is properly governed. According to proper government, a word-internal empty nucleus may not be phonetically realized when it is adjacent to another nucleus which itself is not licensed. If proper government fails, the word-internal empty nucleus must be phonetically manifested. With the phonological ECP taking care of the interpretation of empty nuclei, many so-called syncope phenomena in different languages can be satisfactorily explained (see e.g. Moroccan Arabic discussed in Kaye et al, 1990; French & Yawelmani in Kaye, 1995, etc.)

1.3.5 Tone theory

In the set of elements mentioned above, there are two tonal elements—H and L. As Kaye (1999b) proposed, there are maximally three levels of tonal contrasts on a single nuclear position: H (high tone), L (low tone), and _ (toneless or so-called "mid tone"). With these three levels, human languages can be divided into two types of tone systems, namely, simple systems, which have H and _, and complex systems, which have H, L, and _. On the basis of these two tone systems, Kaye (2000b) further proposed a "complete typology for tonal systems", which is shown below.

(15) A Classification of Tonal Inventories

i. Simple: H or toneless

ii. Complex

i) H, L or toneless.

ii) H, L, HL or toneless.

The simple tone system proposed in (15) can be considered as two-tone system. Logically, there could be three two-tone systems. They could be H and L, H and nothing, or L and nothing. However, (15) shows that the two-tone system can only include H and toneless. To eliminate the other two possibilities, Kaye (2000b) proposed the following principles.

(16) i. L is available for tonal marking only if H is available.

ii. All tonal inventories must contain a toneless expression.

The tonal contrasts then combine with each other to form tonal patterns of the languages.[2] As all the logical possibilities of tones have never been found in human languages, it is claimed that there must be tonal licensing constraints defining the tonal patterns of the languages. Kaye (1999b) thus proposed three Universal Principles of Tone Systems for templatic languages.

(17) Universal Principles of the Tone Systems of Templatic Languages

i. Extended OCP: a given tone may appear only once in a tonal pattern. Thus there cannot be two Hs or Ls in a single tonal pattern, but _ _ _ is well formed.

ii. Every tonal pattern has a head which may be defined positionally or intrinsically.

Let's take Mandarin as an example. There are three tonal contrasts in Mandarin on a single nuclear position. Thus it is categorised as a complex tone system. However, the tonal contrasts do not and cannot combine randomly. From its four tonal patterns, namely, H_, _H, LH and HL, four representations, or tonal patterns to be more exact, are derived—High Level (High spreads to the "toneless" position), High Rising, Low Rising and Falling. The licensing constraint of the tone system of Mandarin can also be derived from the patterns, that is, H is an intrinsic head (Kaye, 1999b).

The above is a sketch of the core notions of GP. Some of the other concepts will be introduced along with the relevant discussion in respective chapters. Next I will show how the dissertation is organized.

[2] Tonal patterns are often confused with tonal expressions, which are usually called "tones". For

1.4 Outline of the Dissertation

The dissertation is divided into six chapters. In the present chapter, i.e. Chapter One, I have briefly introduced some basic notions of GP including the element theory, theories of constituent structure and transconstituent structure, governing relations, phonological ECP and the tone theory.

In Chapter Two, I will discuss the phonological template of HT. It will be shown that HT is a templatic language as the Han languages studied so far, and the structure of the phonological template of HT will be discussed. I will make the claim that the basic phonological template of HT is $O_1N_1O_2N_2$, on the basis of which an augmented template, $O_1N_1O_aN_aO_2N_2$, can be derived. This chapter will also propose the constraints of the augmented template in HT.

The purpose of Chapter Three is to discover the nuclear system of HT, licensing constraints, and phonological representations for the lexical nuclei. I will first discuss the notion of licensing constraints and demonstrate how they work to generate all and only the nuclear expressions found in a given language. Then I will present the licensing constraints in HT, by the way of which the nuclear system of HT will be established. The discussion of the nuclear system of HT will

---

example, Mandarin is always said to have four tones, which are in fact four tonal patterns derived

be further divided into the discussion of the licensing power of elements and spreading operation.

After establishing the nuclear system of HT, Chapter Four aims to discover the onset system of HT and the phonological representations of the onsets. I will first discuss the previous studies on the consonants of two varieties of To, which are spoken in the same county as HT and are thus supposed to be closely related to HT. Then I will focus on the onset system of HT. Constraints with different elements will also be proposed in the process of establishing the onset system of HT.

Chapter Five will concentrate on the interaction across constituents. There are two major types of interaction between constituents in HT. One is interaction between the two nuclei. The other is interaction between onset and nucleus. The latter is further divided into the interaction between the first pair of onset and nucleus and the interaction between N_1 and O_2. The interaction between nuclei consists of spreading, while the interaction between onset and nucleus involves both spreading and specific constraints.

Chapter Six will present the discussion on the tone system of HT. In the first part of the chapter, it will be shown that HT has a complex tone system, i.e. there can be three tonal contrasts on a single nuclear position, which are H, L, and _. In the

---

from the two tonal expressions and one toneless expression of the tone system of the language.

second part, I will discuss the licensing constraints of tones in HT, which include the tonal inventory, spreading, the licensing constraint and the head of the tonal patterns. Comparisons will be made among the tonal systems of Mandarin, Cantonese and HT for the sake of illustration. It will be shown that HT has more similarities with Cantonese in the tonal system.

Finally, I present the major conclusions of this dissertation.

Chapter Two

The Phonological Template[1]

2.0 Introduction

In this chapter, it will be shown that HT is a templatic language as the Han languages studied so far, and the structure of the phonological template of HT will be discussed. A templatic language, as the name suggests, has a template of definite structure for all its minimal phonological domains. This may be understood in contrast to non-templatic languages, such as English, French, etc., which have no fixed number of constituents for their phonological domains and thus can theoretically run indefinitely long. Thus whether a language is templatic or non-templatic is parameterized, which can be characterized as follows.

1 The first draft of this chapter was accepted by WECOL 2000 held in Fresno, California State University, U.S.A. in October 2000.

(1) Phonological template parameter

Templatic? [Yes] / [No]

The typical templatic languages studied so far are Semitic languages (Arabic, Hebrew, Tigrinya and Tigre) and Han languages (see e.g. studies of Mandarin by Chiu 1994, Goh 1996, etc.; Yichang by Pan 2000; Harbin by Xu 2000; Kunming by Chen 2001; Cantonese by Lin 2001, etc.[2]). Kaye has speculated that the longest template should be of four pairs of onset and nuclei, while the shortest can have two.

It should be pointed out that the so-called "syllable", which has been used so often as a conventional term, has no theoretical status in GP as there is no evidence and necessity for its existence as an independent and meaningful constituent. Thus the idea that minimal phonological strings in languages, like those of Han, are of single-syllable is not accepted. Coda, for the same reason, does not exist within the framework of GP either. Instead it is called post-nuclear rhymal position to be precise.

Therefore, based on the theory of constituent structure of GP introduced in the first chapter, the constituents of a phonological template are primarily onset and nucleus. In this chapter, I will propose that HT has a basic $O_1N_1O_2N_2$ template,

[2] Studies on the Han languages mentioned in this dissertation refer to these researches. The references will not be repeated in the following.

from which an augmented template of $O_1N_1O_aN_aO_2N_2$ can be derived. Section 2.1 will first discuss the $O_1N_1O_2N_2$ template, starting with illustrations from the Han languages, especially Mandarin, as it has been widely studied and the phonological template is well established. In section 2.2, the augmented template of HT will be discussed. Section 2.3 will summarize the conclusions of this chapter.

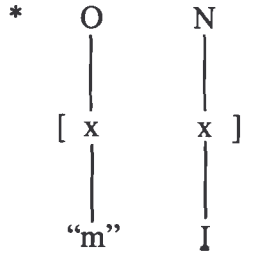
2.1 The $O_1N_1O_2N_2$ Template

It has been mentioned in the above that the basic constituents for a phonological template are onset and nucleus. However, the number of pairs of onset and nucleus varies from language to language. In addition, even if some languages may share the same template, the constraints within the template can be different. Of the languages with $O_1N_1O_2N_2$ template studied so far, the Han languages seem to be the closest to HT in both the template and p-licensing constraints. Thus I will start with the Han languages to illustrate the $O_1N_1O_2N_2$ template.

2.1.1 $O_1N_1O_2N_2$ template of the Han languages

Research on the Han languages done so far has clearly shown a uniform template of $O_1N_1O_2N_2$ structure for all the varieties (see the beginning of the chapter for references). This means that each template must be composed of two onsets and two nuclei. Thus the following two representations for [mi:] 'rice' (2) and [tan] 'talk' (5) are ill-formed in Han languages.

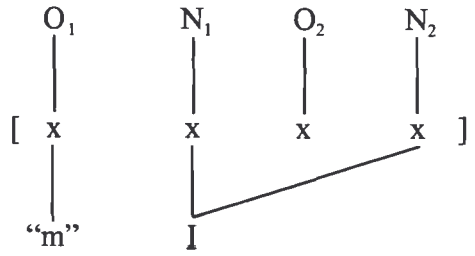
(2) [mi:] 'rice' (tones not included)[3]



The above structure of only one pair of onset and nucleus obviously does not conform to the $O_1N_1O_2N_2$ template. On the other hand, it is contrary to the real length of [mi:] in Mandarin, which in fact occupies two positions instead of just one. The correct representation then should be depicted as follows.

[3] There is an independent chapter discussing the tone systems of Mandarin and HT, thus tones are not included in the analyses of phonological template, onsets and nuclei.

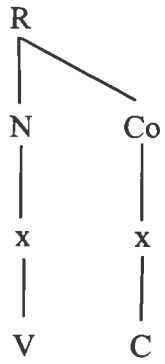
(3) [mi:] 'rice' (tones not included)



In (3), I is the elementary representation for the sound /i/. The length of [mi:] is correctly represented by spreading I in N_1 to N_2. It needs to be mentioned that particles like [lə], [mə], [nə] etc. are not meaningful content words, and are not considered as minimal phonological expressions that take two nuclear positions, and they have no tonal patterns. In addition, they cannot occur independently.

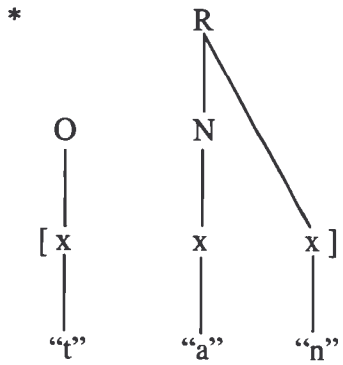
Then what is the structure for words like [tan] 'talk'? The difference between GP's treatment and others' lies in the position of /n/. Traditional generative treatment puts it in the coda position (see e.g. Cheng 1973, Lin 1989, Yin 1989, Duanmu 1990, etc.), as it was generally accepted that a rhyme is itself composed of two sub-constituents: a nucleus (containing a vowel) and a coda (containing post-vocalic consonant) as depicted below.

(4)



Thus, for the structural representation of words like [tan], previous treatments are ill-formed within the framework of GP.

(5) [tan] 'talk'



The reason why those representations are ill-formed is that they violate the principle of Coda Licensing (Kaye 1990:311) in GP which prevents the existence of a post-nuclear rhymal position unless it is followed by an onset position.

(6) Coda Licensing Principle

Post-nuclear rhymal positions must be licensed by a following onset.

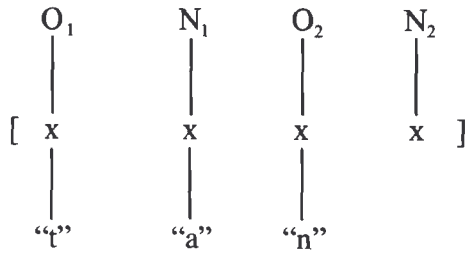
It has been mentioned above that there is no theoretical status for the coda in GP, as there is no evidence for its existence as an independent constituent. However, GP does recognize the post-nuclear position dominated directly by the rhyme, and the principle of Coda Licensing in (6) clearly stipulates that it must be followed by an onset so as to be licensed and governed. The implication of this is that domain-final consonants cannot appear in the post-nuclear rhymal positions. They must be in the onset positions, which in turn must be followed by a nuclear position so as to be licensed. In a word, consonants cannot occur domain-final, as mentioned in Chapter One that there is also a principle of Onset Licensing in GP repeated below for the reader's convenience, even if they are put in the onset position.

(7) Onset Licensing

An onset head position must be licensed by a nuclear position.

Therefore it can only be the nuclei that occupy domain-final position, even if they are empty. Thus the correct analysis for [tan] 'talk' is to put /n/ in the O_2 position instead of the post-nuclear rhymal position proposed in other analyses. The representation is as follows.

(8) [tan] 'talk'



It has been pointed out above that given a certain phonological template, languages may further vary in specific parameterized constraints within the template. Studies on the Han languages have shown that O_2 and N_2 cannot both be filled nor can they be empty, i.e. they cannot both be p-licensed. The P-Licensing Constraint for Beijing Mandarin is as follows.

(9) P-Licensing Constraint (Beijing Mandarin)

N_2 does not contain lexical material.

It needs to be pointed out that the above P-Licensing Constraint is different from the one proposed by Goh (1996a, 1996b). When Goh proposed the P-Licensing Constraint of Beijing Mandarin, He did not realize that N_2 in BJ Mandarin does not contain lexical material. In fact, he put some O_2 in the N_2 position. More and more evidence has shown that /y/ and /w/ in words like [bey] 'cup' and [low] 'building' are in the O_2 position instead of N_2. The underlying constraints are as follows.

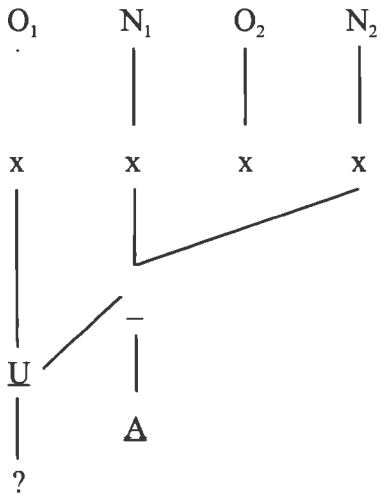
(10) Spreading Constraints (BJ Mandarin)

i. I or U in O_1 spreads to the head of N_1 when O_2 is p-licensed.

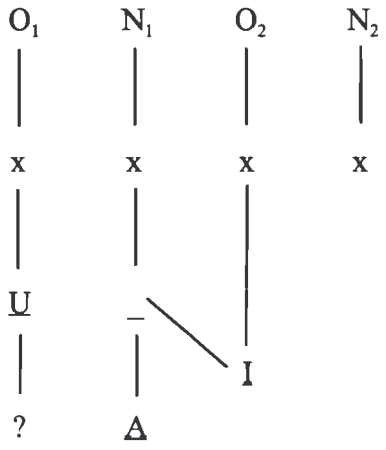
ii. I or U in O_2 spreads to the head of N_1.

To illustrate the above constraints, let us compare [bo] 'wave' and [bey] 'cup'.

(11) a. [bo] 'wave'



b. [bey] 'cup'



In (11a), the O_2 position is p-licensed, thus the U in the O_1 spreads to the head of N_1, achieving a nuclear expression ({A}, U). Then N_1, which is headed now, spreads to N_2. On the other hand, in the second structure, O_2 is not p-licensed and contains ({}, I). In this case, the U in O_1 cannot spread to N_1. Following (10ii), I in O_2 spreads to N_1, achieving ({A}, I).

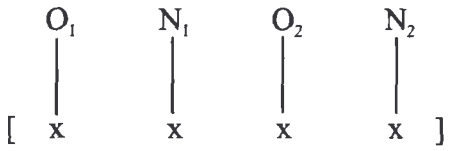
As the study of the Han languages is not my main concern here, more details will not be provided here. Relevant studies mentioned at the beginning of this chapter can be consulted. Having given an illustration of the phonological template with the Han languages, I will now explore the phonological template of HT and its differences and similarities with other templatic languages.

2.1.2 $O_1N_1O_2N_2$ for HT

Although Zhuang has been studied for decades and massive amounts of data were collected during extensive fieldwork, no discussion on the phonological template of Zhuang, both in general and of the varieties spoken in different areas, has ever been presented. Now with the theory of GP, which has been widely applied in the studies of templatic languages, the phonological template of HT can be studied and analyzed. In the following subsections, I will show that the basic and primary

template of HT is a domain of four positions, i.e. two pairs of onsets and nuclei, as shown in (12) below.

(12)



HT shares not only the $O_1N_1O_2N_2$ template with Han languages but also the P-Licensing Constraint and Nucleus Spreading Constraint.

(13) P-Licensing Constraint (HT)

N_2 does not contain lexical material.

(14) Nucleus Spreading Constraint (HT)

N_1 spreads to N_2 when O_2 is p-licensed.

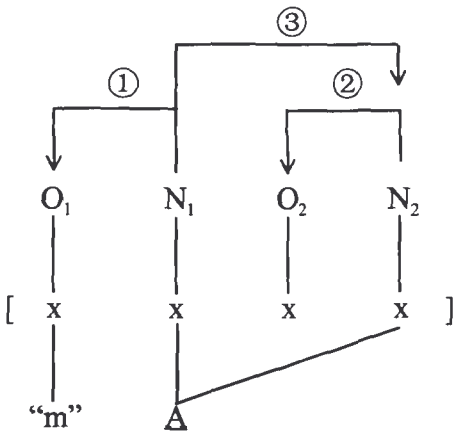
Several conclusions can be derived from the above two constraints. First, N_2 must always be identical with N_1, as it is lexically empty and can be filled only with the material spread from N_1. This eliminates illegal formations like *[tai], [pea], etc., where the two nuclei are different. But words like [ma:], [bi:] are good in HT. The second conclusion is that N_2 and O_2 cannot both be p-licensed, nor can they be both filled. The constraint in (14) has stated that when O_2 is empty, N_2 must be filled with the expression spread from N_1. This implies that N_2 stays empty only when O_2 is filled and blocks the spreading from N_1 to N_2. This supports the

existence of words like [dam] 'pestle', [mən] 'round', etc. and rules out illegal formations like *[lomo]. The following two subsections are to present analyses and evidence for the $O_1N_1O_2N_2$ template, the P-Licensing Constraint and the Nucleus Spreading Constraint of HT.

2.1.2.1 Typical configurations

First let us consider the structure of [ma:] 'dog' depicted below.

(15) [ma:] 'dog'

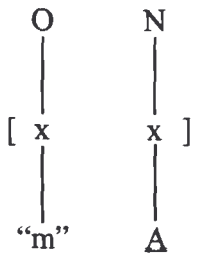


In the above representation, /m/ takes the O_1 position and is followed by N_1 which can and must license the presence of its preceding onset, O_1. The A in N_1 position is GP's elementary representation for /a/. It is clear in the above structure that O_2

is p-licensed, so it is empty. According to the Nucleus Spreading Constraint proposed above, A[4] in N_1 should spread to N_2, occupying one more position. ③ shows a governing relation between N_1 and N_2 at the nuclear projection level. N_2 licenses and governs the position of O_2 as shown in ②, though it is empty.

It may be observed that the $O_1N_1O_2N_2$ template and the P-Licensing Constraint of HT eliminate the existence of words with only one pair of onset and nucleus, as depicted below.

(16) *

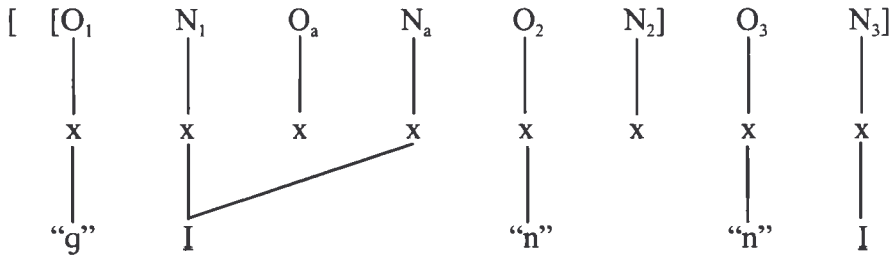


The above structure is ill-formed in HT and cannot be accepted as an independent and meaningful word in the language. However, besides the structure in (12), we do find a few forms with only one onset and one nucleus. They are a very small number of grammatical suffixes that do not form a domain by themselves and do

[4] It is written in elements because the elementary representations for vowels are quite established in GP studies (see e.g. Korean by Kim 1996; Chamorro by Backley & Takahashi 1998; Turkish by Charette & Göksel 1996; to name only a few). However, details about the vowel and consonant system of HT and their elementary representations will be discussed in the following chapters.

not spread. They are always attached to a phonological domain, e.g. [ni] as in [gi:n ni] 'here' and [ða:n ni] 'there' (The structure for words like [gi:n] and [ða:n] will be discussed in Section 2.2 of this chapter). The structure is represented below.

(17) [gi:n ni] 'here'

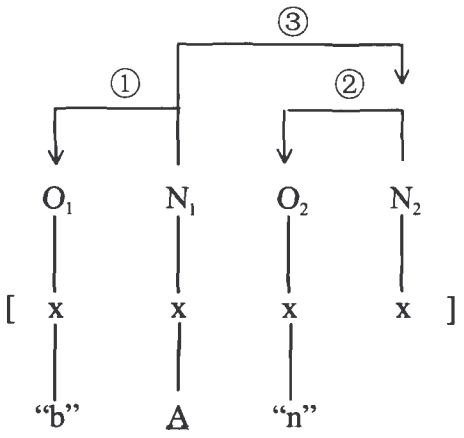


The above structure shows that [ni] is analytic[5] because it cannot share the same domain with [gi:n], otherwise there would be three sets of onsets and nuclei in the minimal domain, which would violate the phonological template proposed above. On the other hand, it cannot be an individual domain either, because it has only one onset and one nucleus, which does not satisfy the condition of $O_1N_1O_2N_2$ as a phonological template of HT. Thus it can only be a suffix that has only one set of onset and nucleus. In fact, suffixes, like [ni], in HT can occur only when they attach to an independent domain and form a morphologically complex structure as (17).

[5] By "analytic" I mean that it is part of a morphologically complex structure and is attached to another phonological domain.

Next let us look at an example with O_2 filled, [ban] 'to divide'. The structure is as follows.

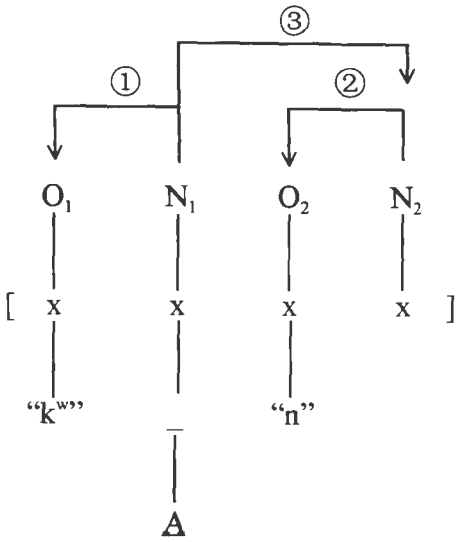
(18) [ban] 'to divide'



As in (15), N_1 licenses its preceding onset, O_1, and licenses and governs the following nucleus, N_2 (process ① and ③respectively). With the reason stated earlier in this chapter—the coda has no theoretical status in GP and a post-nuclear rhymal position must be licensed by a following onset, which itself, in turn, must be licensed by a following nucleus, /n/ is put in O_2 position. As O_2 is not empty, it blocks the spreading from N_1 to N_2. Thus, N_2 stays empty.

Let us now turn to the structure of [k^{w}ən] 'hair' / 'to feed' presented overleaf.

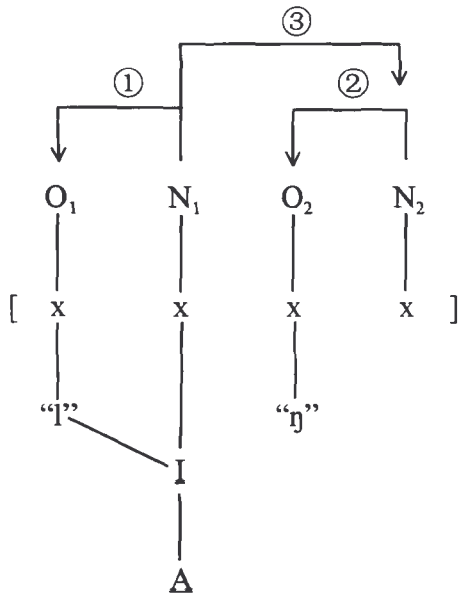
(19) [k^{w}ən] 'hair' / 'to feed'



In structure (19), O_1 is taken by a labio-velar, /k^{w}/. "({A}, _)" in N_1 is the elementary representation for sound /ə/, with A in the operator position and the head being empty. As in the above representations, N_1 licenses its preceding position, O_1, and governs the following nucleus. As it is not p-licensed and contains /n/, O_2 blocks the spreading of material from N_1 to N_2. Thus, N_2 is p-licensed, but licenses and governs O_2. In fact, Chapter Three will show that headless expressions in HT cannot spread anyway, even when the following onset is empty.

Next let us consider the structure for [l^{y}eŋ] 'strength'.

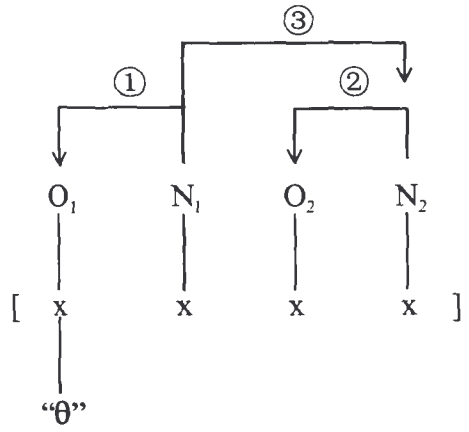
(20) [lʸeŋ] 'strength'



In this structure, the skeletal point dominated by N_1 is associated with a complex nuclear expression, ({A}, I). The following chapters on vowels and consonants will show how the inter-constituent spreading in (20) takes place—the element I in N_1 spreads to the operator position of O_1, achieving /lʸ/. N_1 licenses the preceding O_1 containing /l/ as shown by ①. As /ŋ/, the domain-final consonant, appears in O_2, it cannot be p-licensed, and it prevents N_1 from spreading to N_2. ③ shows that N_2 receives licensing power from N_1 so as to license and govern O_2 in ②.

The analyses presented above have one thing in common: N_1 is never empty, though some of the O_2 positions are empty. Now let us see the structure of a domain with an empty N_1, [θɨ:] 'tiger'.

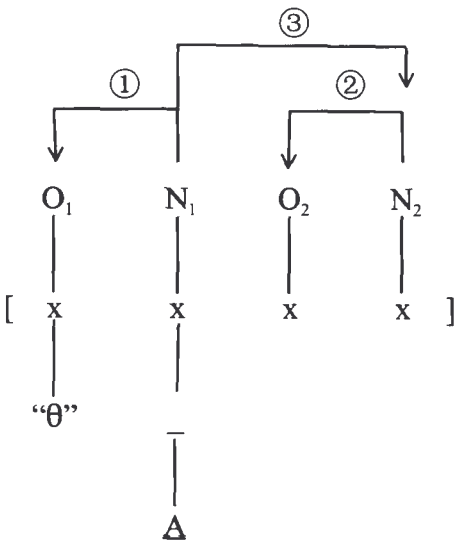
(21) [θɨː] 'tiger'



The above structure shows that /ɨː/ takes two positions. But this is not due to the spreading of the element in N_1 to N_2, as, in this case, N_1 is empty and has nothing to spread. This makes N_2 remain empty. On the other hand, although both O_2 and N_2 are empty in the above structure, they cannot both be p-licensed. Otherwise, one would hear [θɨ] instead of [θɨː]. Thus, in (21), either O_2 or N_2 must not be p-licensed. I propose that O_2 is p-licensed in (21). Thus, N_2 is not p-licensed and has phonetic interpretation, achieving the same empty sound /ɨ/ as in N_1.

Next let us look at the structure of [θəɣ] 'to buy', where N_2 is p-licensed and the empty O_2 is phonetically realized.

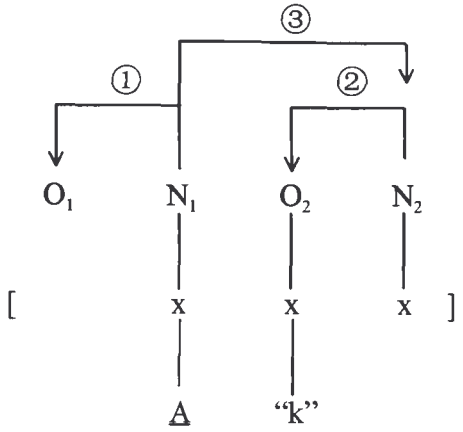
(22) [θə̂ɣ] 'to buy'



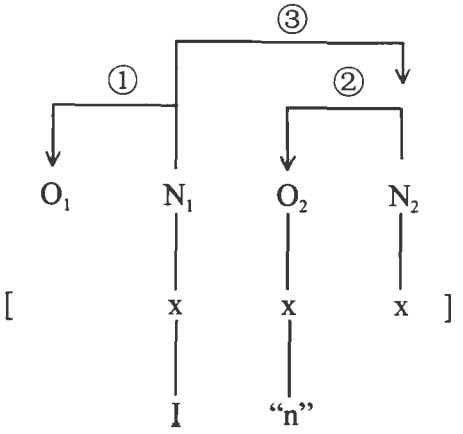
In (22), the ({A}, _) in N_1 first licenses O_1 position containing /θ/. Then, unlike the N_1 in (21) which is empty, N_1 in (22) has an element A in it. However, it is in the operator position with the head being empty. As headless expressions in HT may not spread (see Chapter Three for detailed discussion), the material in N_1 cannot spread to N_2. So N_2 remains empty and is thus p-licensed. As N_2 is p-licensed, O_2 cannot be p-licensed, because, theoretically, at least one of the pair of onset and nucleus must not be p-licensed and heard. Otherwise, there would be no way to know the existence of a pair of onset and nucleus. Thus O_2 in (22) is phonetically heard, resulting in an empty expression, /γ /.

Finally let us look at a structure where O_1 is empty. (23) below are two examples.

(23) a. [ak] 'chest'



b. [in] 'cigarette'



The above two examples are different from the others provided earlier in that they each have a lexically empty O_1. I propose that in these cases, not only is O_1 lexically empty. It has no position as well, because according to The I-Spreading Constraint in HT, which will be presented in the following chapters, I in N_1 should spread to O_1 when O_2 is filled. However, in (23b), no such spreading takes place,

implying that O_1 has no lexical position. More examples will be provided in Chapter Three and Chapter Four to prove this. Besides the lexically empty O_1, the governing and licensing relationships in (23) are the same as in earlier examples.

From the above examples and their phonological structures, it is not difficult to make the following observations: (i) N_1 is always the one to initiate licensing and governing—it licenses a preceding O_1 and a following N_2 which may then license O_2, yet it is never licensed or governed. (ii) The material for O_2 is very limited. In fact, a closer look at the O_2 position shows that in HT, O_2 can only dominate /m/, /n/, /ŋ/, /w/, /y/, and clipped consonants like /p/, /t/ and /k/. As for N_2, it can only be filled with elements spread from N_1 when O_2 is p-licensed and does not block the spreading. (iii) There is one and only one skeletal point dominated by each onset and nucleus. (iv) O_1 and O_2 cannot both be empty.

The first observation shows that N_1 is the head of the domain. This follows from the Phonological Licensing Principle proposed by Kaye (1990):

(24) Phonological Licensing Principle

Within a domain, all phonological units must be licensed save one, the head of that domain.

This principle entitles N_1 to be the head of the phonological domain in the phonological template of HT so as to license and govern other constituents of the

domain. In addition, Harris (1992, 1997) proposed another universal principle of GP: Licensing Inheritance, which provides an explanation as to why O_2 can only dominate a limited number of elements. The principle is as follows (Harris, 1997: 340).

(25) Licensing Inheritance

A licensed position inherits its a-licensing potential from its licensor.

Adapting Goldsmith's notion of autosegmental licensing, Harris used the term a-licensing "to refer to the sanctioning of a subsegmental units of melody by the syllabic positions to which they are attached." (Harris, 1997: 337) And by a-licensing potential, he means "the position's ability either (i) to directly a-license a melodic unit or (ii) to confer a-licensing potential on another position". (Harris, 1997: 340) He further states that a licensed position is weaker than its licensor—"Its potential is diluted as a result of being acquired from another position, namely the licensor..." (Harris, 1997: 340) This gives us a reason to expect N_2 to be "weaker" than N_1 in terms of its licensing properties. Thus, it can be concluded that in HT, N_1 is the head of the domain and is unlicensed at all projections. O_2, on the other hand, as is licensed by a weak position N_2, cannot be as 'strong' as O_1 either. It must dominate a fewer number of consonants than O_1. Finally, as there is one and only one skeletal point associated to each onset and nucleus, it can be concluded in HT that, neither nucleus nor onset branches.

To sum up, five conclusions can be drawn based on the above observations and investigation of the ONON template of HT: (i) The minimal domain in HT consists of exactly two nuclei; (ii) If N_2 is p-licensed then O_2 is not p-licensed; (iii) N_1 is the head of the domain; (iv) Constituents in HT, both onset and nucleus, do not branch; (v) O_1 and O_2 cannot both be empty.

2.1.2.2 Additional evidence

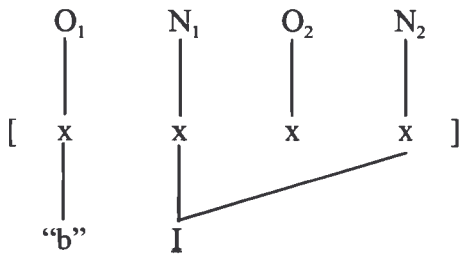
The reasoning and conclusions proposed above are based mainly on data from HT proper. They are supported by comparative evidence from Jiangnan village. What has been found is that /i:/ in the Jiangnan dialect corresponds to /əy/ in HT, when there is no following consonant. The correspondence is systematic and consistent. Some examples are listed overleaf.

(26) Correspondence of /i:/ in Jiangnan To and /əy/ in HT

Jiangnan To	HT	Glossary
[vi:]	[vəy]	"comb"
[bi:]	[bəy]	"fat"
[ni:]	[ɗəy][6]	"gut"
[ti:]	[təy]	"dense"

Structures for [bi:] and [bəy] are presented below.

(27) a. [bi:] 'fat' (in Jiangnan To)



[6] /ɓ/ and /ɗ/ in HT correspond to /m/ and /n/ in Jiangnan To.

b. [bəy] 'fat' (in HT)

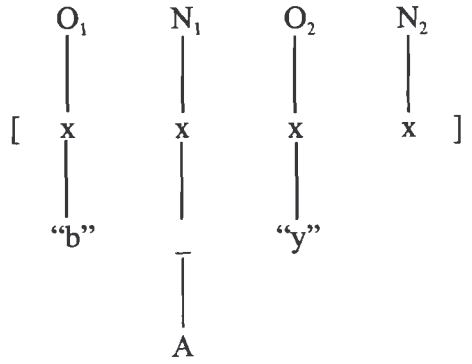
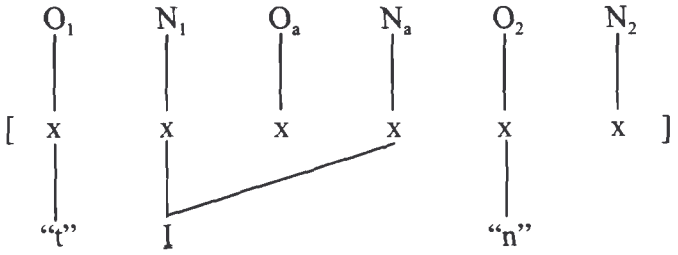


Figure (27a) shows that I in N_1 spreads to N_2 as O_2 is empty and does not block the spreading. In (27b) for [bəy], however, as O_2 position is occupied by I, which blocks spreading between N_1 and N_2, and also because N_1 is occupied by an empty-headed expression that cannot spread, N_2 remains empty.

The story changes when O_2 is occupied in the Jiangnan dialect—the change of /i:/ in Jiangnan To to /əy/ in HT never happens. So the word [ti:n] 'to know' in Jiangnan dialect does not correspond to *[təyn] following the general pattern. Rather it is also [ti:n] in HT. The explanation for this difference in correspondence provides a strong piece of evidence for the phonological template of HT proposed above. As I have argued earlier that the phonological template in HT is $O_1N_1O_2N_2$ and there are no branching onsets or nuclei. A representation like *[təyn] would be ill-formed, as /y/ occupies the O_2 position and there is no more room for /n/. Thus, instead of being *[təyn] following the above pattern, [ti:n] in Jiangnan

dialect is also [ti:n] in HT, the structure for [ti:n] is shown below[7].

(28) [ti:n]



So far it has been shown that HT shares some similarities with the Han languages, namely, the $O_1N_1O_2N_2$ template, non-branching constituents, N_1 as the head of the domain, P-Licensing Constraint and Nucleus Spreading Constraint. Be that as it may, an important difference between HT and the Han languages is that the $O_1N_1O_2N_2$ template in HT has a variant—$O_1N_1O_aN_aO_2N_2$, which is called the augmented template. The following section is to introduce the notion "augmented template" first, then discuss its behavior in HT.

2.2 $O_1N_1O_aN_aO_2N_2$—Augmented $O_1N_1O_2N_2$

The preceding section introduced the typical phonological template of HT, which

[7] The following section will discuss the status of O_aN_a, which is part of the augmented template.

contains exactly two sets of onsets and nuclei. This pattern has been found in some languages, especially in the Han languages. However, this is not the complete story for the phonological template of HT. Data in this language show that although in both Han languages, like Mandarin, and HT, there is spreading between two nuclei, the distribution of this nuclear spreading in the two languages is different. In Mandarin, for example, N_1 spreads to N_2 only when O_2 is p-licensed. It implies that in Mandarin, when both the two onsets are filled, it is impossible to get a long vowel. There is no position for it, as the template is strictly $O_1N_1O_2N_2$. In HT, however, we do find an onset preceded by a long vowel that takes two nuclear positions in a single domain. I suggest that this is because in HT, an augmented template, $O_1N_1O_aN_aO_2N_2$, exists.

2.2.1 Augmented template in general

The notion of augmented template was first proposed by Lowenstamm in the studies of Semitic languages[8]. In Semitic languages, e.g. Arabic, Hebrew, Tigrinya and Tigre, major categories, like noun, verb, etc., do not contain vowels. Instead

[8] It is through personal contact that the notion "augmented template' was discussed. Jean Lowenstamm and Mohaut Guerssel have discussions and manuscripts on the augmented template of Semitic languages. But there is no previous publication to refer to.

they are just skeletons, e.g. "ktb" 'to write' in Arabic[9]. The augmented template is dictated or activated by morphology or morphosyntax of the language.

2.2.2 Augmented template in HT

In HT, it is found that the O_a and N_a positions contain no lexical material. As a consequence, N_a only receives phonological material from the preceding nucleus, N_1, while O_a is always empty. The major features for the augmented template in HT are proposed as follows.

(29) Constraints of the Augmented Template (HT) (*I*)

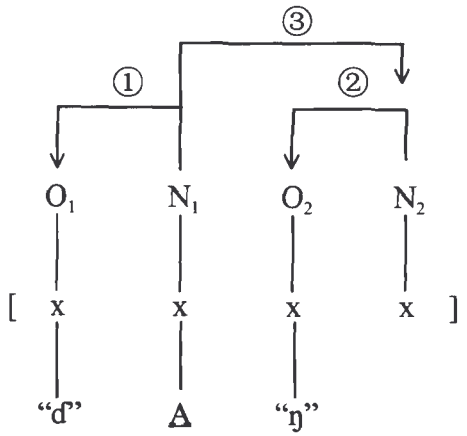
i. No lexical material is associated with the augmented set of onset and nucleus, i.e. O_a and N_a;

ii. N_1 spreads to N_a.

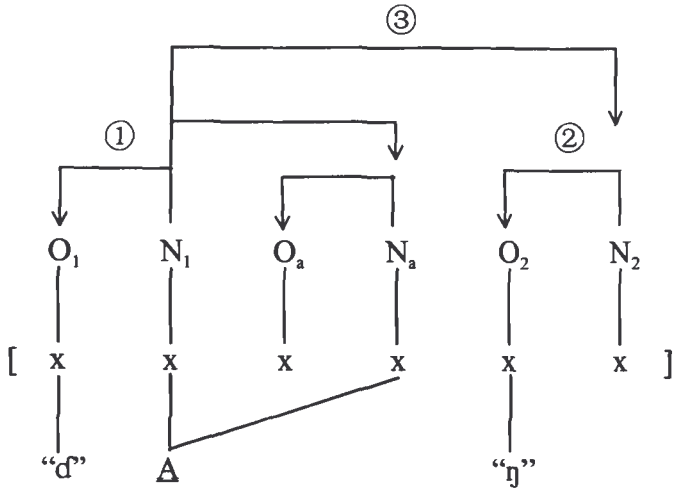
From this constraint, it is not difficult to deduce that no new tones are created at the augmented positions, as the tonal elements H and L are lexical material as well. The following is an illustration of the difference between the two templates with the structure of word [ɗaŋ] 'nose' and the one of [ɗa:ŋ] 'body / health'.

[9] Data of these languages are from the lectures and seminars by Jonathan Kaye at Guangdong University of Foreign Studies.

(30) a. [ɗaŋ] 'nose'



b. [ɗa:ŋ] 'body/health'



(30a) and (30b) clearly show that [ɗaŋ] and [ɗa:ŋ] have different phonological

templates. In structure (30a), there is no extended set of nucleus and onset. N_1 is filled with the element A, resulting in a short /a/. In (30b), on the other hand, there is an augmented pair of nucleus and onset. N_a receives spreading from N_1, achieving a long /a:/.

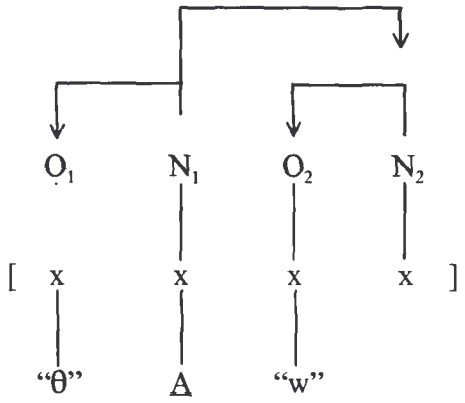
As it has been proposed in (29) that O_a and N_a have no lexical material and N_a always receives spreading from N_1, we can derive that in the augmented template of HT, O_2 may not be p-licensed, while N_2 must always be p-licensed. Otherwise there would be three nuclei phonetically interpreted in a phonological domain. This can be summarized as another constraint of the augmented template of HT presented below.

(31) Constraints of the Augmented Template (HT) (II)

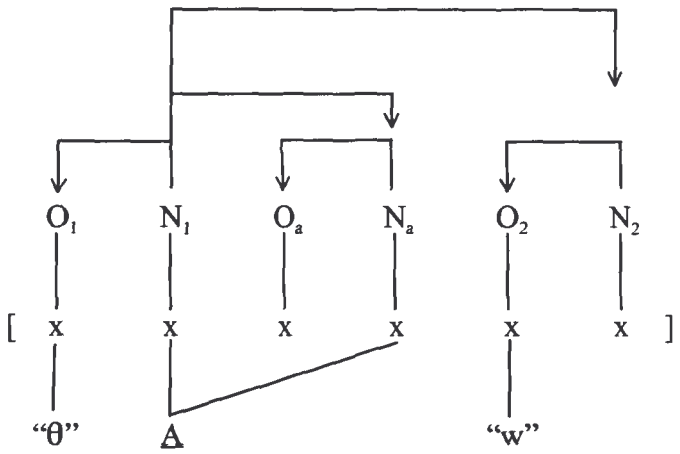
In the augmented template, O_2 can never be p-licensed while N_2 must always be p-licensed.

The licensing and governing relation within the augmented template is the same as that of the $O_1N_1O_2N_2$ template, namely N_a gets the licensing power from N_1 to license O_a, as shown in the above two representations. Below is one more comparison between two words of different phonological templates, i.e. $O_1N_1O_2N_2$ vs. $O_1N_1O_aN_aO_2N_2$.

(32) a. [θaw] 'harvest'



b. [θa:w] 'pretty'



2.3 Summary

In this chapter, I have discussed the phonological template of HT and made a comparison with that of the Han languages, especially Mandarin. The phonological template of HT is typically $O_1N_1O_2N_2$, from which an augmented template, $O_1N_1O_aN_aO_2N_2$, can be derived. However, the augmented pair of O_aN_a contains no lexical material. As a consequence, N_a only receives phonological material from the preceding nucleus, N_1, while O_a is always empty. Along with the principles of the phonological structure within a domain in HT, the following conclusions about the templatic structure of HT can be drawn:

1. Phonological Template of HT

   HT typically has an $O_1N_1O_2N_2$ template, from which $O_1N_1O_aN_aO_2N_2$ can be derived.

2. Constituent Structure (HT)

   In HT, constituents do not branch.

3. P-Licensing Constraint (HT)

   N_2 cannot contain lexical material.

4. Nucleus Spreading Constraint (HT)

   N_1 spreads to N_2 when O_2 is p-licensed.

5. Head of the Domain (HT)

   N_1 is always the head of the domain.

6. Onset Constraint (HT)

   O_1 and O_2 cannot both be empty.

7. Constraints of the Augmented Template (HT)

   i. No lexical material is associated with the augmented set of onset and nucleus, i.e. O_a and N_a;

   ii. N_1 spreads to N_a.

   iii. In the augmented template, O_2 can never be p-licensed while N_2 must always be p-licensed.

The above seven principles show that HT is very similar to the Han languages with respect to the phonological template, the main difference being that HT has an augmented template.

Chapter Three

The Nuclear System

3.0 Introduction

Systematic transcriptions on varieties of the Zhuang language have been carried out in the past decades. This has laid a solid foundation for further theoretical studies to some extent. Nevertheless, these are not phonological studies, but merely list all the possible speech sounds of the language. Licensing constraints enable us to work out the complete phonological system of a language—how to get a well-formed phonological expression and why others are ill-formed. In this chapter, we will first discuss licensing constraints in general and their roles. Then by discovering the specific licensing constraints for the nuclear expressions of HT at the lexical level, the nuclear system of the language will be established.

The chapter is structured as follows: in 3.1, the notion of licensing constraints

proposed by Kaye (2000a), Cobb (1993), and Charette and Göksel (1994), and examples from some languages will be discussed. We will see what licensing constraints are and what they do for the representations of nuclear expressions. Then in 3.2, the nuclear system of HT will be presented as well as the representations of the nuclear expressions in this language. There are two subsections in 3.2. In 3.2.1 I will discuss the licensing power of three elements A, I and U in HT. I will discuss what they can license in a nuclear expression. In 3.2.2, we will see the spreading behavior of elements in the nuclear position in HT—what can spread, when they can spread, etc. The last section, 3.3, will summarize the discussion in the whole chapter.

3.1 Licensing Constraints

The representational system for phonological expressions in GP differs from that of other theories in many ways. One of these differences is that the ultimate constituent of an expression is not the phonological feature but the element. Thus all phonological expressions—nuclear expressions, onset expressions, etc., are made up either of an element or combinations of elements. As this chapter concerns

only the nuclear system of the language in question, the discussion will be limited only to elements A, I and U,.[1]

The original motivation of proposing the notion of licensing constraints is to explain restrictions on the combinatorial properties of elements. The underlying assumption here is that "any syntactically well-formed combination of elements should be present in a phonological system unless explicitly excluded" (Kaye, 2000a). As the studies carried out so far show, no single language can exhaust all the theoretically possible combinations of elements. There must be some language-specific rules regulating the behavior of these elements. Licensing constraints were proposed to define the lexical set of phonological expressions in a given linguistic system.

As this chapter mainly concerns A, I and U, I will list all the possible combinations of the three elements without any constraints and then illustrate the principle of licensing constraints by using data from some languages. Following the definition of a phonological expression provided in the first chapter, which requires that each expression must be composed of a head and operator(s), there are 20 possible

[1] These three elements are also used in the representation of onsets. Next chapter will show how they work to form an onset system of the language, together with other elements in the representational system of phonological expressions in GP.

expressions involving the elements A, I and U. Twelve headed expressions are generated as in (1) below and eight headless expressions as shown in (2).

(1) ({}, I) ({}, U) ({}, A)

({A}, I) ({I}, A) ({I}, U) ({U}, I) ({U}, A) ({A}, U)

({A, U}, I) ({I, U}, A) ({A, I}, U)

As earlier chapters have illustrated, the operator can be a set in which the order of elements is not relevant. Thus in the above table, a phonological expression like ({I, U}, A) is equivalent to ({U, I}, A). What is important in the operator part or a phonological expression is the presence or absence of a given element. This is also true for headless expressions.

(2) ({A}, _) ({U}, _) ({I}, _)

({U, I}, _) ({A, I}, _) ({U, A}, _)

({A, U, I}, _) ({ }, _)

The expressions in (2) share one common property—all are headless.

The two tables, (1) and (2), have shown the 20 possible structures from the free

combination of elements A, I and U. However, as has been stated earlier in this chapter, no studies have found all 20 expressions within a single phonological system. This is due to the central role of licensing constraints—setting rules for the presence and absence of phonological expressions in a given linguistic system. The mechanism of licensing constraints thus reduces the twenty possibilities to all and only the phonological expressions allowed in the given language.

Licensing constraints define the way in which elements can combine at the level of lexical structure. This involves what elements may be heads and whether heads can license operators. For example, the licensing constraints for branching nuclei in English (Kaye, 2000a) require that all expressions are headed and the element I can only be a head. So headless expressions like ({A}, _), ({U, I}, _), etc. are excluded as long vowels. Whenever I is in a phonological expression, it must be in the head position, ({}, I) and ({A}, I). On the other hand, elements A and U can both be head and operator because the licensing constraints do not prohibit this. Thus expressions like ({A}, U) as in "sew" and ({U}, A) as in "saw" are acceptable in English nuclei.

In addition to the licensing constraints of the elements' properties of being heads or operators, another constraint is whether heads can license operators, as a licensing

relation is supposed to exist between head and operator of a phonological expression. It has been noted that in many languages the element "A" cannot license operators. Thus in these languages, ({}, A) is well-formed, whereas all expressions with the form (X, A), where X represents a non-null set of elements, are unacceptable. Take Turkish as an example[2]. In Turkish, there are ({}, A) and ({A}, I), while ({I}, A) is ruled out. In other words, when A is in the head position, the set of operators must be empty, as A is not a licensor according to the licensing constraints of Turkish. It has no power to give permit to the presence of other elements in its operator part.

Besides the above two constraints, there can also be constraints on the combination of elements. For instance, in some languages, like English, there is one constraint saying that *I and U may not combine* (Kaye, 2000a). In this case, expressions like ({I}, U), ({U}, I) or ({I, U}, _) will be illicit. From the above brief introduction of licensing constraints, it is not difficult to conclude that once licensing constraints for a given language are correctly set, phonological expressions at the level of lexical representations of the language will be appropriately generated, as we will be able to test the legitimacy of phonological expressions and eliminate those that do not conform to the licensing constraints. Moreover, in addition to providing

[2] See Charette & Göksel (1998) for details.

good explanations to phonological events in a language, we will be able to predict various phonological phenomena in that language.

Having introduced the notion of licensing constraints and its central role, I will now propose the specific licensing constraints placed upon the lexical representations of nuclear expressions in HT, and show how the proposed licensing constraints generate the nuclear system of HT as well as explain some phonological phenomena of the language.

3.2 The Nuclear System of HT

Although a massive amount of fieldwork and transcriptions have been carried out on the Zhuang language in past decades, no systematic studies of the nuclear system of the language at the lexical level have been carried out. The closest varieties of To to the one in question that have been transcribed are those spoken in the same county, Tiandeng. As neither descriptive nor theoretical studies on HT are available, I myself have collected the raw data. The following analyses of the nuclear system of the language are based on these data. There will also be

comparisons with other known varieties of To spoken in Tiandeng County for explanation and illustration.

Data from the language show the following set of surface vowels:

(3) Surface vowels in HT:

{a, i, u, o, ie, ə}

However, these are not all the surface vowels. There is another similar set of vowels shown in (4)[3].

(4) {a:, i:, u:, o:, e:, ɤ:}

(3) and (4) are listed in the following table (5), together with examples, for the reader's convenience.

[3] Besides the above two sets of surface vowels, there are also some problematic speech sounds in HT, e.g. /ia/, /iu/, /iə/, /ua/, /ui/, /uə/. It seems that they could be understood either as light diphthongs or as an onset plus a nucleus. But their restricted distribution implies that /i/ and /u/ in these speech sounds are not part of the light diphthongs, but rather of the onsets. Detailed discussions on this will be provided in the next chapter of the onset system of HT.

(5)

Vowel	Example	Glossary
a:	ɗa:ŋ	'body'
a	ɗaŋ	'nose'
i:	ti:n	'to know'
i	lin	'tongue'
u:	tu:	'head'
u	kum	'pit'
e:	de:	'he/she'
ie	biet	'eight'
o:	ho:p	'box'
o	ɠop	'frog'
ə	dzəɣ	'yes/be'
ɨ:	θɨ:	'tiger

At the first sight, it seems that (4) is just the longer counterpart of those in (3). However, /ə/ in (3) and /ɨ:/ in (4) have no counterparts in the other set. Nor do /e:/ and /ie/. On the other hand, as all the surface long vowels in (4) show up only when O_2 is empty or when there is an augmented position in the template, we may well assume that there are only seven nuclei in the nuclear system of HT as follows:

(6) {a, i, u, o, e, ə, ɨ}.

This inventory may seem problematic as /e/, instead of /ie/, is included. We will

leave this to the discussion of spreading, because there is evidence proving that /ie/ is derived from /e/.

Now that we have proposed the seven nuclei of HT, licensing constraints should be established to show how they are derived and why other expressions are absent from the above nuclear system. To do so, we need first to define the L-structure expressions. As in syntax, L-structure is the level of lexical representations, which contains phonological representations before they undergo any phonological events. After applying the phonological rules, P-structure can then be derived from L-structure. The L-structure expressions in a language are generated by the licensing constraints of the language. In order to discover the complete licensing constraints of HT, I will first discuss the licensing constraints with the three elements—A, I and U in the following subsection. Constraints on spreading will be discussed in the next subsection.

3.2.1 Licensing power of A, I, and U

In this subsection I will discuss the licensing constraints involving elements A, I

and U. As has been claimed earlier in this chapter, the assumption is that the three elements are supposed to have the same licensing power if there is no constraint restricting them from doing so. However, the discussion provided above implied that in no linguistic system can the three elements be totally equal in terms of licensing power. Particularly, it has been found that in many languages, "A" turns out not to be a licensor (Kaye, 2000a). So a provisional licensing constraint of HT is proposed in (7), with which we can then reduce the 20 possible expressions in (1) and (2) to the seven expressions representing those surface vowels in (6).

(7) Licensing Constraint in the Nuclear System of HT

I and U must be heads.

Constraint (7) eliminates all the expressions where I and U are operators. They are crossed out in the following table, where all the 20 possible nuclear expressions of (1) and (2) are put together.

(8) Expressions remaining after the application of (7)

({ }, A)	({A}, I)	~~({A, I}, U)~~	({A}, _)	~~({A, I, U}, _)~~
({ }), I)	({A}, U)	~~({A, U}, I)~~	~~({I}, _)~~	({ }, _)
({ }, U)	~~({I}, A)~~	~~({I, U}, A)~~	~~({U}, _)~~	
	~~({I}, U)~~		~~({A, I}, _)~~	
	~~({U}, A)~~		~~({A, U}, _)~~	
	~~({U}, I)~~		~~({I, U}, _)~~	

The above table clearly shows that constraint (7) has eliminated all the illicit expressions in HT, leaving the following seven expressions that are supposed to represent the vowels in (6).

(9) The L-structure phonological expressions in HT

({ }, A)	({ }), I)	({ }, U)
	({A}, I)	({A}, U)
({A}, _)	({ }, _)	

Note that from (7) other two facts in the nuclear system of HT can be derived. First, I and U cannot combine in a phonological expression as they must be heads and GP requires that there can be only one head in an expression. Secondly, as (7) has required that I and U must be heads and cannot be in the operator position, it can be

derived that no head can license them as operators. Besides I and U, there is only the element A that can be a head. Thus, it can be derived that A does not license I or U as operators.

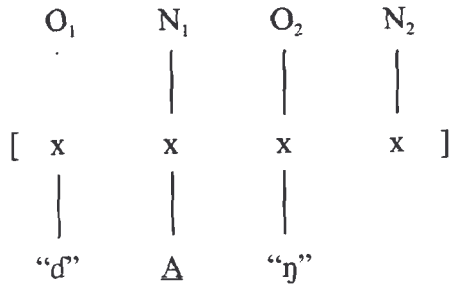
Through the above means of applying licensing constraints in (7), there are exactly seven expressions left, which are given in (9). It should be pointed out that the licensing constraints allow for the "empty" expression, ({ }, _), shown in the last column in (9). The representation for each nuclear expression found in HT is thus proposed as follows:

(10) Representations of nuclear expressions in HT

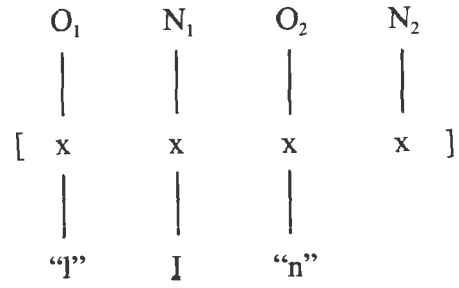
GP Expression	Letter Representation
({ }, A)	a
({ }, I)	i
({ }, U)	u
({A}, U)	o
({A}, I)	e
({A}, _)	ə
({ }, _)	ɨ

The following examples illustrate the grammatical forms proposed in (9) and (10).

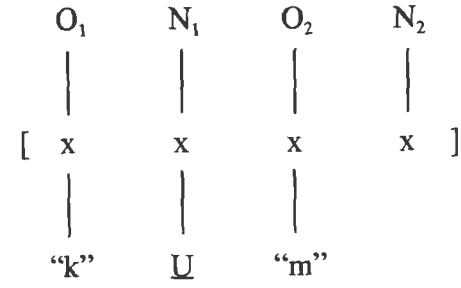
(11) a. [ɗaŋ] 'nose'



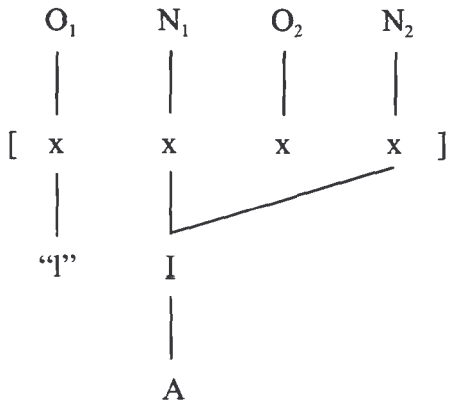
b. [lin] 'tongue'



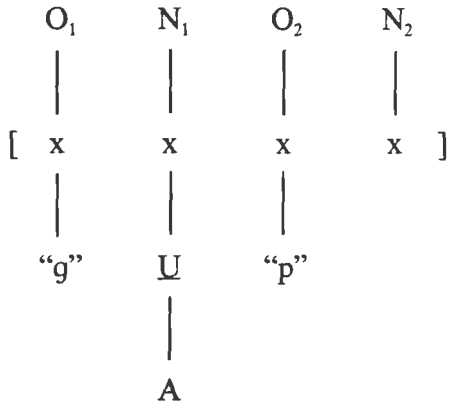
c. [kŭm] 'pit'



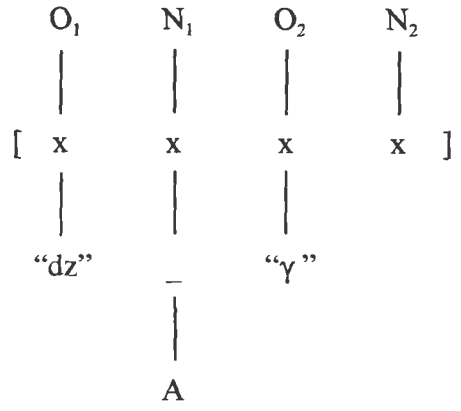
d. [le:] 'look'



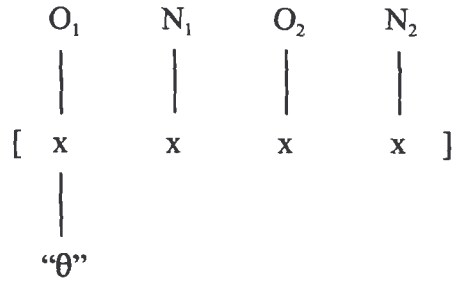
e. [go^{p}] 'frog'



f. [dzəɣ] 'yes/to be'



g. [θɨː] 'tiger'



In the above configurations, the first five examples, i.e. structures (a) to (e), are easy to understand. The first three consist of a simplex nuclear expression, ({}, A), ({}, I) and ({}, U) respectively, while each of the last two contains a complex expression. Furthermore, the five expressions undergo the same process in deriving their respective long versions. This will be shown in the next subsection on spreading.

Examples (f) and (g), however, are different from the first five in that they are headless. It might seem that there is also a length distinction between (f) and (g). It would be wrong to consider /ɨ:/ to be a long vowel derived by the spreading of N_1 to N_2. As explained in 2.1.2.1, the fact that here /ɨ:/ seems to take two nuclear positions is not because of the spreading of elements. The phonological expression for /ɨ/ is actually empty, which means it has nothing to spread. The reason for the long /ɨ:/ is that N_2 is completely empty and cannot be p-licensed. So N_2 must have a phonetic realization as an un-p-licensed position. This realization is identical to N_1. This gives the impression of a long vowel. In fact, in the next subsection we will show that /ə/ is always short in HT, whereas /ɨ/ always appears to be long. This is due to the constraints on spreading in the language.

There is also no example of a short /e/ given above. This is because, in the surface vowel inventory of HT, /e/ is always long. It only appears when O_2 is empty, in which case it will consistently spread to the following nucleus, N_2, achieving a long /e:/, e.g., [ɓe:] 'sheep'. Then does this mean that the nucleus /e/ cannot be followed by a filled O_2? The answer is that it can appear in a phonological domain where O_2 is not p-licensed, but the head of the nuclear expression will spread leftward, to O_1. This is part of what I will discuss in the next subsection.

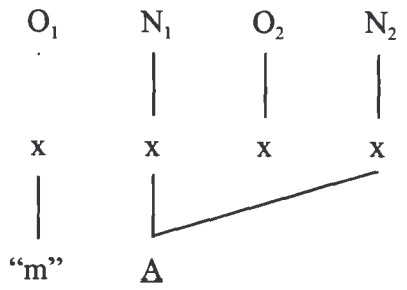
3.2.2 Spreading

The last subsection has provided the licensing constraints for elements A, I and U. They have explained the presence and absence of phonological expressions at the lexical level. However, this is not the complete picture of the nuclear system of the language. We must explain why some nuclear expressions, namely ({ }, A) for /a/, ({ }), I) for /i/, ({ }, U) for /u/, {A}, U) for /o/ and ({A}, I) for /e/, can be extended to the following empty nuclear position, while some, i.e. ({A}, _) for /ə/, cannot; why some, e.g. ({ }, _) for /ɨ/, must always be realized as a long vowel; and why the head of the nuclear expression for the nucleus /e/ must always spread to the preceding onset when O_2 is not p-licensed. This subsection will provide answers for these remaining problems in the nuclear system of HT. It will show that constraints on spreading are the key to these problems. The following is divided into two parts. In the first part I will talk about conditions for spreading and non-spreading, and the second part will focus on spreading in the augmented template.

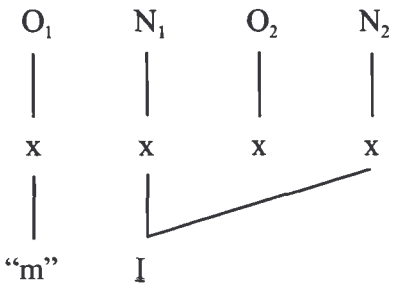
3.2.2.1 Spreading and non-spreading

The last section has shown the typical distributions of the lexical nuclear expressions of HT, and has provided licensing constraints to explain why other expressions are not allowed in the phonological system of the language. However, one major problem remains to be solved: why the vowels /a/, /i/, /u/, /o/ and /e/ may spread, while the vowel /ə/ cannot. The structures are given below.

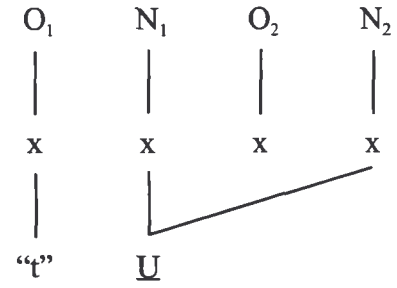
(12) a. [ma:] 'dog'



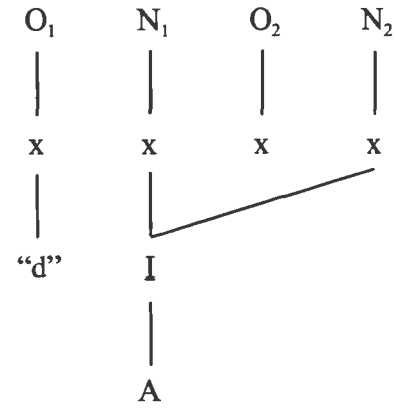
b. [mi:] ‘mom’



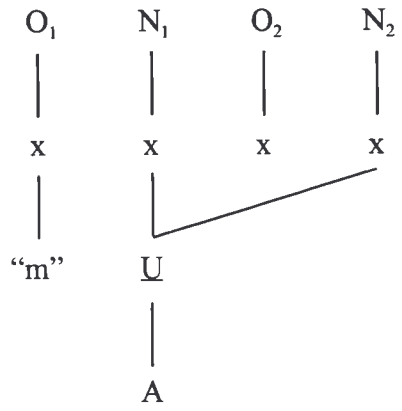
c. [tu:] ‘head’



d. [de:] ‘he/she’

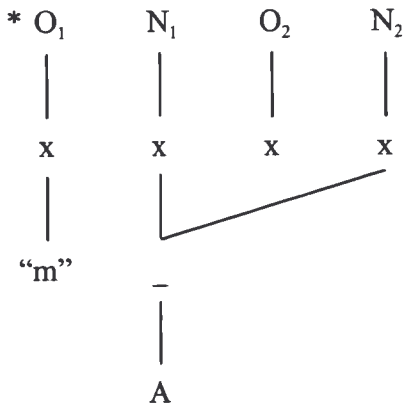


e. [mo:] 'ox'



The above structures show that /a/, /i/, /u/, /o/ and /e/ may spread to N_2 when O_2 is empty. Such spreading cannot be observed for /ə/. Thus, no structure like the one shown below is acceptable.

(13) *[mə:]



A comparison between the nuclear expressions of (12) and that of (13) shows that the major difference between the two is that the expressions in (12) are headed while that of (13) is headless. Accordingly, I propose the following constraint on spreading.

(14) Spreading Constraint in the Nuclear System of HT

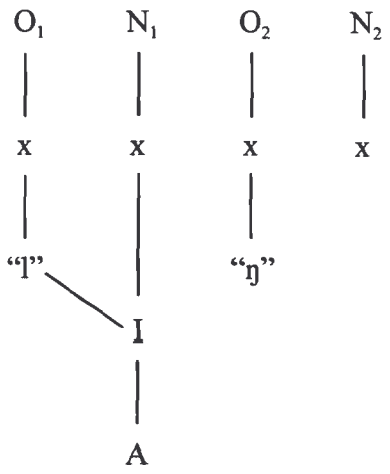
Only Headed expressions may spread.

With this additional constraint, we can then explain why /ə/ does not have a long version in the set of surface vowels.

Till now, we have been discussing nuclear spreading to the right. In fact, in HT, spreading from right to the left occurs more frequently than from left to right. Let us look at one kind of right-to-left spreading in the nuclear system of HT. The next chapters will provide more examples of this kind of spreading in the onset and tone system of the language. The right-to-left spreading is the answer to why /ie/ is derived from /e/. A comparison of HT with two other varieties of To spoken in the same county shows that /ie/ cannot be found in these two varieties. They have /e:/ instead of /ie/. So words like [kien] 'arm', [ɗiet] 'sunshine', [kiek] 'guest / visitor',

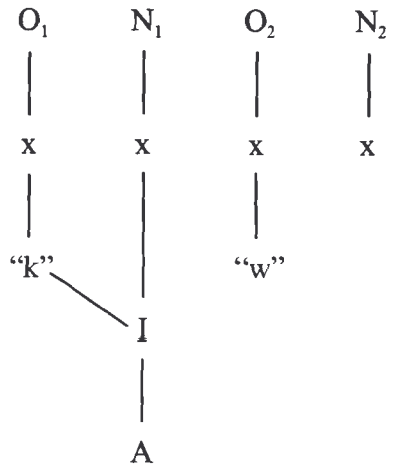
etc. in HT correspond to [ke:n], [ne:t], and [ke:k] in the other two varieties (see e.g. Qin, 1996)[4]. This gives us a hint that /ie/ is derived from its lexical form /e/, with the head of the phonological expression spreading to O_1. So /ie/ is in fact /ʸe/. The derivation is illustrated in the following examples.

(15) a. [lʸeŋ] 'strenght'



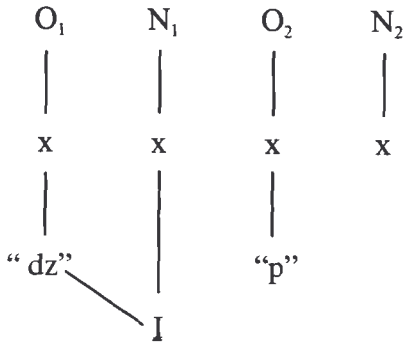
[4] Tones and onset difference are ignored here.

b. [k^{y}ew] 'ball'

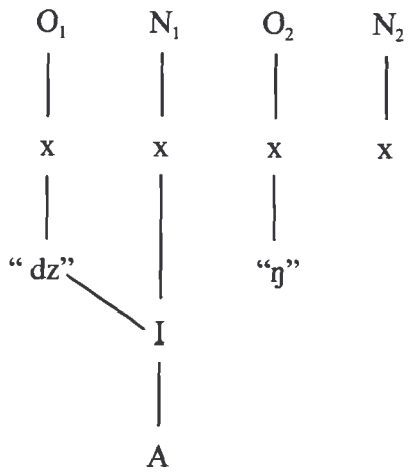


The above two structures show the process of I-spreading from ({A}, I) to O_1 when O_2 is occupied. More evidence of the I-spreading shows up when we compare the distribution of some surface onsets in HT. As will be discussed in the next chapter, some surface onsets are derived due to I-spreading. The following are examples showing how the lexical onsets /dz/ and /s/ derive into /j/ and /x/ when N_1 contains I and O_2 is filled.

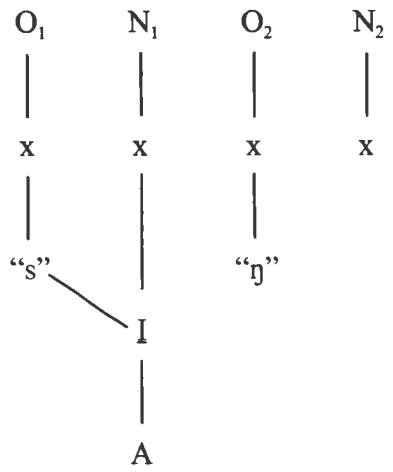
(16) a. [jip] "to hurt"



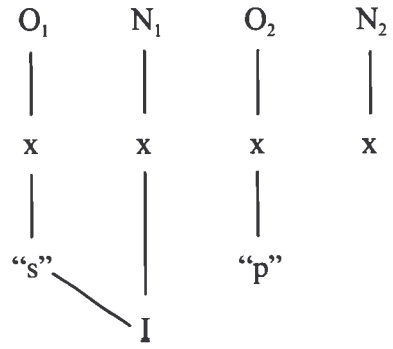
b. [jeŋ] "to snap"



c. [xeŋ] "try one's best"



d. [xip] "to smell"

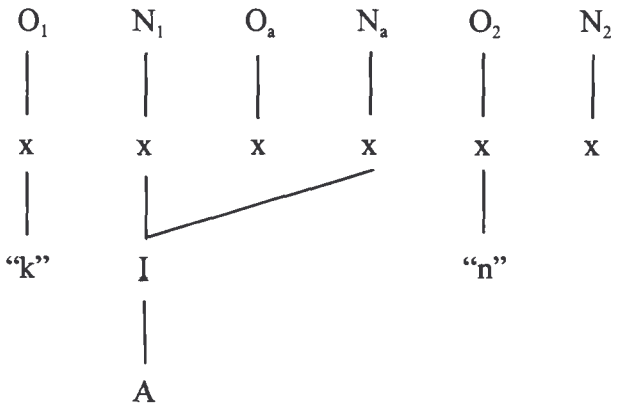


As onsets will be discussed in the next chapter, detailed representations for the onsets above will be provided then.

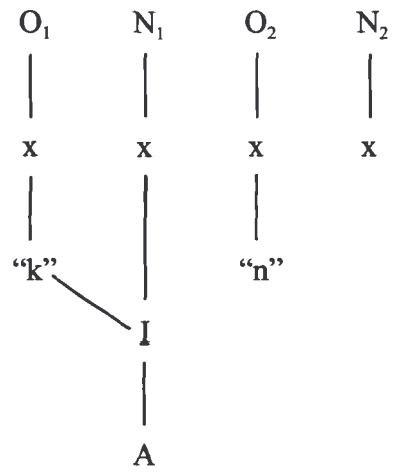
Another important fact in the correspondence between /e:/ and /ʸe/ is the asymmetry

of length. In the above examples of [ke:n], [ne:t], and [ke:k] in Jiangnan To, /e:/ is always long. In other words, it always takes two positions. The counterparts in HT, however, are always short. The structures for [ke:n] in Jiangnan To and [kʸen] in HT are shown as follows.

(17) a. [ke:n]

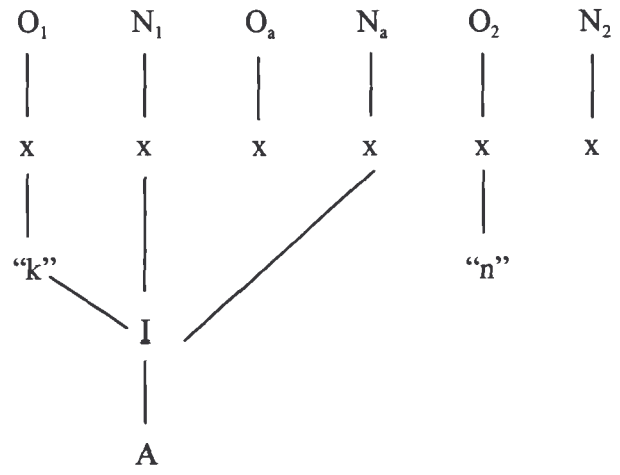


b. [kʸen]



It is obvious that there is spreading in both of the structures. But the direction is different. In (17a), the spreading is from left to right. N_1 spreads to N_a. In (17b), the spreading goes from right to left. The element I in N_1 spreads to O_1. From this spreading behavior, we may conclude that in the nuclear system of HT, I must spread. If it cannot spread to the right, it will spread to the left, and vice versa. However, it cannot spread to both the right and the left at the same time. Thus, expressions like *[k^{y}e:n] are eliminated in HT. Its structure would be as follows.

(18) *[kye:n]



I thus propose the following I-Spreading Constraint in HT.

(19) I-Spreading Constraint in the Nuclear System of HT

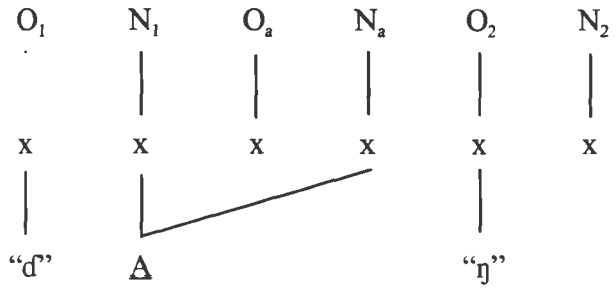
I in N_1 must spread. If it cannot spread to the right, it must spread to the left; and vice versa.

The above constraint provides explanation to the length asymmetry of /e:/ and /ᵛe/ both within HT and among To dialects.

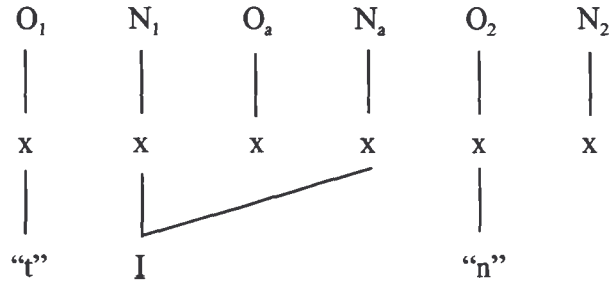
3.2.2.2 Spreading to the augmented position

The structures shown in the above are mostly in the basic phonological template of HT, i.e. $O_1N_1O_2N_2$. In this part, I will provide some structures from the augmented template and illustrate how the nuclei spread or do not spread to the augmented position, N_a. The following are augmented templates, which show how nuclei spread to the augmented position.

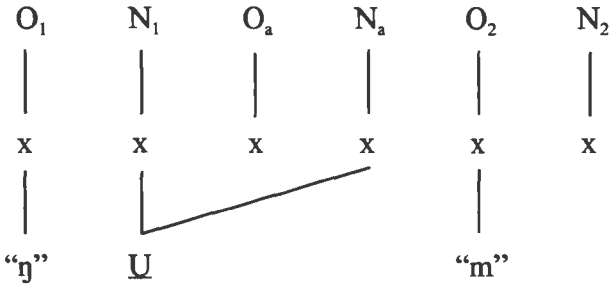
(20) a. [ɗa:ŋ] 'body/health'



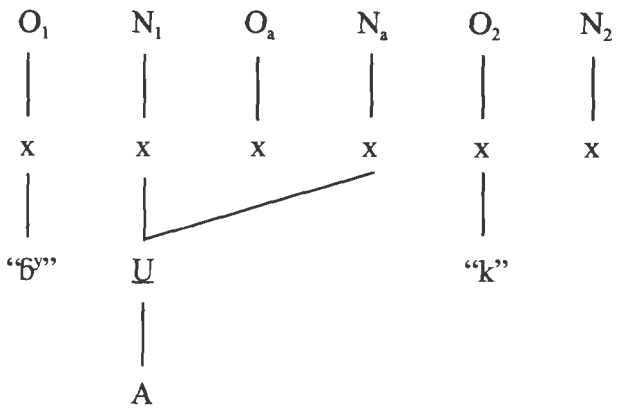
b. [ti:n] 'to know'



c. [ŋu:m] 'cave'



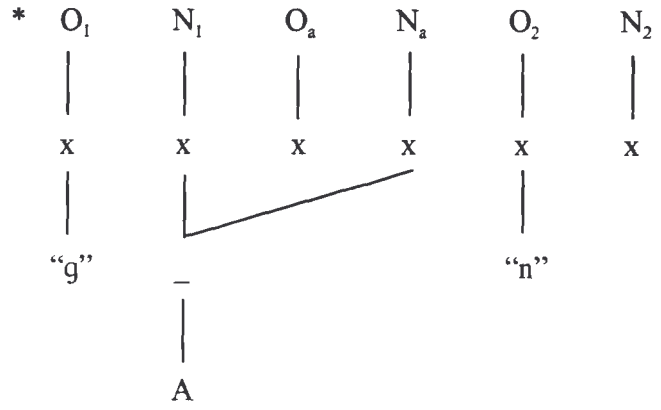
d. [ɓʸo:ᵏ] 'flower'



The above structures have shown how /a/, /i/, /u/ and /o/ spread in the augmented template. As they are all headed expressions, there is no problem for them to spread to the following nuclear position, which is N_a in this case. Because of constraints (19), the nucleus /e/ cannot appear in the augmented template, even though it is also a headed expression.

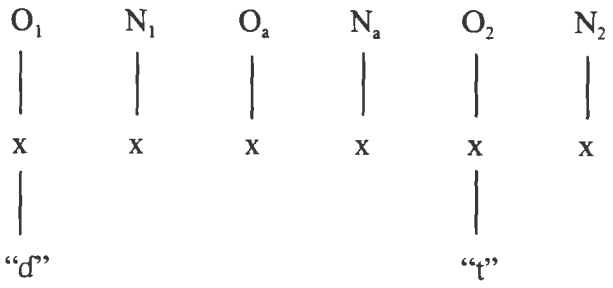
On the other hand, /ə/ must not spread being a headless expression. So there are no words like *[gə:n], the structure of which would be as follows:

(21) *[gəːn]



The above structure would violate the spreading constraint proposed in (14). However, we do find /ɨː/ followed by consonants /n/, /ŋ/, /p/, /t/ and /k/. This might be posed as a counterexample of the behavior of /ɨ/ proposed in the above: it cannot spread to N_2 as it is a totally empty expression. Therefore N_2 should be phonetically realized. Nevertheless, this is not the case with /ɨ/ in the augmented template and words like [dɨːˀ] 'hot', [mɨːŋ] 'brook', etc. are not counterexamples to the above discussion of the behavior of /ɨ/ and the constraint proposed in (14). For a clear illustration, let us see the following structure first.

(22) [dɬːˈ] 'hot'



The above structure shows the situation of /ɬ/ in the augmented template. As N_1 is empty, it has nothing to spread. Thus it stays at its original place, N_1. As a result, the augmented position N_a in this case remains empty, as there is no spreading from N_1. This makes N_1 and N_a share the same property of being empty expressions. Thus in this case, although N_1 cannot spread to the augmented position, N_1 and N_a are not p-licensed. It has been proposed in the second chapter that the extended positions in the augmented template have no lexical material, i.e. they are empty. As for the case in HT, N_a can only get elements from N_1, while O_a is always empty. Thus the above structure is not rejected in the phonological system of the language. The two empty positions, N_1 and N_a are then phonetically realized in the augmented template. Then one might further ask why N_2 is p-licensed? In fact, it is not /ɬ/ that p-licensed N_2. It is because that if N_2 were also not p-licensed, there would be three nuclei realized in one phonological template. This would violate the constraint of the phonological template of the language. Chapter Five will provide discussion of why the O_2

position following /ɨ/ can only be filled by a restricted number of consonants.

The situation of /ə/ in the augmented template of HT has been said to be different from that of /ɨ/, though they are both headless expressions. /ɨ/ is absolutely empty while /ə/ still has an element "A" as the operator. Although /ɨ/ cannot spread to the augmented nucleus, N_a is also empty. Their co-existence is not in conflict with the augmented template and its constraints. The situation with /ə/, on the other hand, is different. As /ə/ is headless and cannot spread to its augmented nuclear position, N_a remains empty. The non-uniformity of their elementary property violates the constraints for the augmented template of HT. Therefore structure (22) is not acceptable in the phonological system of HT, where N_1 and N_a do not share the same material.

The strange behavior of /ə/ and /ɨ/ might well have been puzzling to descriptive linguists of Zhuang language. They either transcribe /ə/ as /ɨ/ and list some exceptions (see e.g. 壮语方言土语音系) or just remove it (see e.g. Qin, 1996:249).

3.3 Summary

In this chapter, I first introduced the notion of licensing constraints and how they operate in phonological systems. Next I considered the nuclear system of HT and its licensing constraints. In discovering the licensing constraints of the language, the discussion was divided into two parts. One was about the licensing constraints with respect to the three elements A, I and U. This is used to determine the nuclear expressions at the lexical level. I then discussed the constraints of spreading. Putting the discussion of the two parts together, the licensing constraints of the nuclear system of HT can be summarized as follows:

1. Licensing Constraint in the Nuclear System of HT

   I and U must be heads.

2. Spreading Constraint in the Nuclear System of HT

   Only Headed expressions may spread.

3. I-Spreading Constraint in the Nuclear System of HT

   I in N_1 must spread. If it cannot spread to the right, it must spread to the left; and vice versa.

The following chapters will provide more discussion and evidence for these constraints so as to confirm the explanation to respective phenomena.

Chapter Four

The System of Onsets and Their Representations

4.0 Introduction

In last chapter, we have considered the nuclear system of HT and its licensing constraints. In this chapter we will focus on the onset system of the language and present the underlying representations for these onsets. In the discussion of the phonological representations for nuclei, three of the six elements, namely A, I and U, were introduced. In defining the representations for onsets, the role of the other three elements, H, L and ?, will be shown.

As in the process of discovering the nuclear system of HT, the discussion of the representations of onsets is also based on constraints. This may involve constraints between onset and nucleus as well as constraints between two onsets. These inter-constituent interactions will be systematically discussed in the next chapter, while this chapter will just mention a few used for analyzing the onset system. In addition to presenting the onset system of HT, the chapter will also provide some

explanations for related phenomena.

The chapter is divided into three sections. In section 4.1, earlier descriptive studies on the consonant inventory of two varieties of To spoken in the same county as HT will be provided. It will be shown that the main disagreement concerns the position of some constituents—whether they should be in the onset or the nucleus position. In section 4.2, I will discuss the onset system of HT. In this core part of the chapter, the apparent onsets of HT and a brief discussion on the distribution of some onsets will be provided. These raw data will be used for the further analyses in the following subsections. It will be shown that some of the surface onsets are derived from underlying lexical onsets through spreading from the licensing nucleus. In subsection 4.2.2, I will discuss the onset system of HT and present the phonological representations for the lexical onsets through systematic analyses of two constraints—the U-Constraint and the Empty Nucleus Constraint. In 4.2.2.1, I will discuss the I-based onsets in HT. Then in 4.2.2.2, U-based onsets of HT will be discussed. Constraints involving the empty nucleus will be discussed in 4.2.2.3. Finally, a summary of this chapter will be given in section 4.3.

4.1 Previous Studies of the Consonants of Some Varieties of To

Although no work has been carried out on HT and thus no data nor studies about the

onset system or consonantal inventory of this language have appeared, there are many systematic transcriptions of the other varieties of To (see e.g. Qin 1996, etc.). This includes some of the varieties spoken in the same county, i.e. Tiandeng County. In this section, we will look at the earlier analyses of the consonant inventory of just two varieties of To. They are both spoken in Tiandeng County and are thus supposed to be closely related to HT. The descriptive studies of the two varieties of To show two different sets of surface consonants. They agree to a large extent, but seem to have different opinions about the position of some sounds. The following are the two sets of consonants proposed for the two varieties of To spoken in Tiandeng County.

(1) Surface consonants of Jiangnan To (Qin, 1996:250)

p	ph	m	f	v		
t	th	n	ɬ	l	ts	
		ɲ	x	j		
k	kh	ŋ	kj	kw	?	h

(2) Surface consonants of Lichuan To (Guangxiqu Yuwei Yanjiushi, 1994:364)

p	p'	m	f	w		
t	t'	n	l	θ	ts	s
j						
k	k'	ŋ	?	h		
pj	p'j	mj	kj	vj		
kw	k'w	ŋw				

Although they use different symbols, it seems that they generally agree on the 'simplex' consonants, and set (2) includes most of the consonants listed in set (1).[1] It may be noticed that the two sets have several differences, e.g. the /v/ in Jiangnan To corresponds to /w/ in Lichuan To. This should remind us of a similar situation in Mandarin, in which /v/ and /w/ are two variants of one onset representation and they make no significant difference. In fact, people in the northeast, e.g., prefer to use /v/ instead of /w/ when it is in the O_1 position by itself. Thus, take words [wang] 'net' and [wey] 'for' as examples. Northeasterners will be more likely to pronounce them as [vang] and [vey], while people in southern China will say [wang] and [wey]. Native speakers of Mandarin simply ignore this difference, as it is phonologically insignificant.

[1] For some unknown reason, ɓ and ɗ in HT do not have independent correspondences in these two varieties. Rather they are included in m and n respectively.

The above phenomenon is also found in HT. In the process of data collecting and fieldwork, I found that, when /w/ is in the O_1 position by itself, native speakers of HT ignore the difference between /w/ and /v/ as well. They take both as correct, though they prefer to use /v/. In the following section, I will provide a detailed discussion showing that they are just two variants of one phonological expression and do not form a minimal pair in HT. In addition to the correspondence between /w/ and /v/, the situation with /ɬ/ and /θ/ is the same. More comparisons with other varieties of To (e.g. Longzhou To, Laibin To, etc.) show that, just like /v/ and /w/, /ɬ/ and /θ/ are complementary in different varieties, i.e. they cannot both exist within one variety of To. In addition, they correspond to each other in different varieties. So [θi:m] 'sharp' in Lichuan To is [ɬi:m] in Jiangnan To.[2] Therefore they are not a contrasting minimal pair either. However, further systematic studies on the other two varieties are needed to prove that they have the same representation. Be that as it may, /ɲ/, /x/ in Jiangnan To and /s/ in Lichuan To seem to form contrasts in the two varieties.

In addition to the above 'simplex' consonants, the two studies have also listed some onsets with /j/ and /w/ respectively. It is obvious that the major problem with the above analyses is not the number of onsets they have. It is that they did not provide any evidence to prove that /j/ and /w/ should be in the onset position or the nucleus position. This common and fundamental shortcoming of descriptive studies makes

[2] As the chapter is about the onset system, tone change will not be discussed here.

the above analyses unreliable. Furthermore, the above two analyses did not distinguish the consonant inventory for O_1 position from the inventory for O_2. This may well be due to the convention of Chinese linguistics, which does not do phonological analysis on the basis of onsets and nuclei, but rather onsets and rhymes. They do not recognize that the phonological template for the Han languages is ONON instead of OR[3]. Studies on Zhuang phonology have taken this position for granted and applied it to the phonological analysis of varieties of To. So N_1 and O_2 have been analyzed as a rhyme instead of two independent constituents. Undoubtedly, this has added much redundancy to the system. It is also problematic when there is a 'long vowel' instead of a rhyme composed of a nucleus and an onset, which has been shown in the second chapter. On the other hand, if we distinguish the two sets of consonants for O_1 and O_2 respectively as the phonological template suggests, the whole phonological system of the language will be more concise and less confusing. The constraints will be more transparent as well.

As the first theoretical study of HT, which is spoken in the same county as the above two varieties, the following is an analysis of the onset system of HT and the phonological representations of onsets. Unlike the above descriptive studies, it will not only provide the consonant inventory but also provide evidence for the validity of the inventory.

[3] See Chapter Two for details.

4.2 The Onset System of HT and Their Representations

In the above section, I provided two consonant inventories proposed in previous studies on two varieties of To spoken in Tiandeng County. In this section, I will first provide the raw data of the consonants of HT and their respective distribution. Then I will present the onset system of the language and the phonological representations for these onsets. It will show that some of the surface consonants are lexically in the onset position, while some are derived due to spreading from the licensing nucleus. I will also use the headship of phonological expressions to explain some of the onset behavior.

4.2.1 Surface onsets and their distribution

By doing massive and systematic transcription and fieldwork, the set of contrasting onsets was established. It has been pointed out above that, since there are two onset positions, i.e. O_1 and O_2, there should be two sets of consonants. As has been discussed in Chapter Two, the set of consonants for the second onset is always smaller than the set for the first onset. Or we may say that consonants in the second onset are usually a subset of consonants in the first onset. The two sets of consonants in HT and their examples are shown in (3a) and (3b) below.

(3) a. Surface consonants in the O_1 position of HT

p	b	m	ɓ	f	v		
t	d	n	l/ð	ɗ	θ	dz	s
y		ɲ				j	x
k		ŋ				g	h
p^y	b^y	m^y	$ɓ^y$			g^y	h^y
k^w		$ŋ^w$	l^w	$ɗ^w$	$θ^w$	g^w	s^w

Surface consonants	Examples	Glossary
p	pa:	'quilt'
b	boy	'to fill'
m	ma:	'to come'
ɓ	ɓi:n	'to fly'
f	fay	'fire'
v	vəy	'comb'
t	ta:ŋ	'tail'
d	da:	'river'
n	naŋ	'cover'
θ	θa:y	sand
ɗ	ɗam	'dark' / 'black'
l/ð	ða:	'I' / 'me'
dz	dzaŋ	'scale'
s	sam (si)	'to nag'
y	yak	'to tear'
ɲ	ɲan	'to force'
j	jim	'to taste'
x	xik	'ruler'
k	kɨ:	'eggplant'
g	ga:ŋ	'jaw'
ŋ	ŋa:y	'lunch'
h	hu:	'something'
p^y	p^ya:	'mountain'
b^y	b^ya:	'fish'
m^y	m^ya:	'loose'
$ɓ^y$	$ɓ^y$a:	'to scratch'
g^y	g^yap	'to catch'
h^y	h^ya:	'to threaten'
k^w	k^wən	'to feed'
g^w	g^we:	'to reap'
$ŋ^w$	(van) $ŋ^w$a:	'yesterday'
$θ^w$	$θ^w$a:	'lock'
$ɗ^w$	$ɗ^w$ay	'stair'
l^w	l^wəy	'to flow'
s^w	s^wa:y	'softheaded'

b. Surface consonants in the O_2 position of HT

y

w m p

n t

ŋ k

ɣ

Surface consonants	Examples	Glossary
y	may	'to like'
w	yaw	'worried'
m	ta:m	'to carry'
p	dap	'to pile'
t	bat	'to sweep'
n	hon	'to beat'
k	dək	'to fish'
ŋ	dzɨ:ŋ	'to raise'
ɣ	θəɣ	'to buy'

It seems that the first four rows of onsets in (3a) on page 6 could be the basic inventory of onsets in HT. As for the remaining onsets in (3a), there is an additional component—either /y/ or a /w/. There are two possibilities for the position of /y/ and /w/—the O_1 or N_1 position. Words like [bʸa:] 'fish', [kʷən] 'to feed', etc. could be analyzed either as [bʸa:] and [kʷən], in which 'y' and 'w' are taken as part of the O_1 position, or as [bia:] and [kuən], where they are considered as part of a light diphthong with /a/ and /ə/and so in a nuclear position. However, the distribution of

these consonants or light diphthongs shows that if /ue/, /uə/, /ua/ were light diphthongs, we would have to explain why they have a very limited distribution, i.e. they can be found only after a very limited number of onsets. To be more specific, we have to find the answer to why words like [mia:] 'loose', [ŋua:^{k}] 'nod', [kuey] 'to ride', [gue] 'to reap' are well-formed, while forms like *[tia:], *[kia:], *[nua^{k}] or *[tuey] are impossible. The underlying assumption here is the principle of free distribution. If [ia], [iu], [io], [ua], [uə], [ue], etc. are well-formed light diphthongs in HT, they should be able to appear after any onset unless explicitly and systematically excluded. Yet no such constraints are found in HT. Furthermore, if the above combinations with /i/ and /u/ were light diphthongs, we would have to add more expressions to the nuclear inventory of the language, typically those mentioned in the above, to which neither systematic distribution nor sound licensing constraints would be possible. On the other hand, putting /y/ and /w/ in the onset position poses no such problems. Later discussion on constraints in HT in this chapter will provide more evidence on this.

Besides the issue of the position of /y/ and /w/, there appears to be another problem in the above two sets. It seems that two of the onsets in (4) cannot be found in (3), namely /w/ and /ɣ/. This does not mean that the above claim that onsets in O_2 are a subset of those in O_1 is wrong. I have claimed earlier in this chapter that /v/ and /w/ are just two variants of the same phonological expression. Moreover, native speakers of HT do not distinguish the two and use both of them interchangeably for O_1, though /v/ is preferred. They are in fact the same onset in HT. In the following

discussion, I will use /v/ when in the O_1 position, and use /w/ in the other cases.

As for /γ /, it is the phonetic realization of an empty phonological expression. So lexically, it is the same as the empty O_1 in words like [ak], [on], etc. The difference between them is just whether it is p-licensed or not. The distribution and detailed discussion of the phonological behavior of HT onsets will provide more evidence.

Having had a brief discussion of some apparent problems with the two sets of surface onsets, I am going to provide details of the onset system of HT in the following subsections. As there has been no previous analysis on the onset system of any varieties of To, the following work will be the first, which is based on the six elements A, I, U, H, L and ?.

4.2.2 Lexical onsets and their representations

In the following, we are going to see which of the surface onsets presented in (3) are lexical onsets and which are derived. Representations for the lexical onsets will be provided. First, let us have a look at the tables setting out the general distribution of the surface onsets and surface nuclei in HT. In the tables, examples are provided with the well-formed sequences. The empty spaces indicate accidental gaps. The asterisk "*" indicates ungrammatical sequences.

(4)

	a:	a(:)+O_2	ə:	ə + O_2	u:	u(:)+O_2	o:	o(:)+O_2
	*	a^k	*	ən	*	um	*	op
p	pa:	pat	*	pən		pu:k	po:	poŋ
b	ba:	ba:n	*	bən	bu:	buy	bo:	bop
m	ma:	ma:k	*	mən	mu:	mum	mo:	mo:n
ɓ	ɓa:	ɓa:n	*	ɓəy		ɓuŋ	ɓo:	ɓo:t
f	fa:	faŋ	*	fən		fuk		fow
v	va:	vam	*	vəy	*	*	*	*
t	ta:	tay	*	təy	tu:	tu:m	to	tom
d	da:	daɣ	*	dən	du:	duk	do:	dop
n	na:	nam	*	nəɣ	nu:	nu:k	no:	no:k
θ	θa:	θa:k	*	θən		θu:n		θo:n
ɗ	ɗa:	ɗaŋ	*	ɗəŋ		ɗuŋ	ɗð:	ɗot
l/ð	ða:	ða:k	*	ðən	ðu:	ðu:ŋ	ðo:	ðot
dz	dza:	dzaŋ	*	dzəɣ	dzu:	dzuk	dzo:	dzop
s	sa:	sam	*	sət	su:	suk	so:	so:n
j	*	*	*	*	*	*	*	*
x	*	*	*	*	*	*	*	*
y	ya:	yaŋ	*		yu:	yuk	yo:	yok
ɲ	ɲa:	ɲaŋ	*	ɲəy		ɲup		ɲo:m
k	ka:	kaw	*	kən		kum	kŏ	kom
g	ga:	gay	*	gən	gu:	guk	gŏ:	go:t
ŋ	ŋa:	ŋaw	*	ŋən	ŋu:	ŋu:m	ŋò:	ŋo:k
h	ha:	ham	*	hən	hu:	hun	ho:	hop
p^y	p^ya:	p^yak	*			p^yu:m		p^yoy
b^y	b^ya:	b^yap	*	b^yəɣ	b^yu:	b^yu:y		b^yon
$ɓ^y$	$ɓ^y$a:	$ɓ^y$a:y	*		$ɓ^y$u:			$ɓ^y$o:k
m^y	m^ya:	m^ya:p	*	*	*	*	*	*
g^y	g^ya:	g^yaŋ	*		g^yu:			
h^y	h^ya:	h^yat	*		h^yu:			
k^w	k^wa:	k^wan	*	k^wəy	*	*	*	*
g^w	g^wa:	g^wan	*	g^wən	*	*	*	*
$ŋ^w$	$ŋ^w$a:	$ŋ^w$ak	*	$ŋ^w$ən	*	*	*	*
$θ^w$	$θ^w$a:	$θ^w$a:y	*	$θ^w$ən	*	*	*	*
$ɗ^w$	$ɗ^w$a:	$ɗ^w$ay	*	$ɗ^w$əy	*	*	*	*
l^w	l^wa:	l^wat	*	l^wəy	*	*	*	*
s^w	s^wa:	s^wa:y	*	s^wəy	*	*	*	*

	i:	i(:)+O_2	e:	e + O_2	ie:	ie + O_2	ɨ:	ɨ: + O_2
	*	in	*	eŋ	*	*	*	*
p	pi:	pi:k		*	*	pieŋ	pɨ:	
b	bi:	bit	be:	*	*	biek	bɨ:	bɨ:ŋ
m	mi:	mi:ŋ	me:	*	*	mien	mɨ:	mɨ:ŋ
ɓ		ɓi:n	ɓe:	*	*	ɓiet	ɓɨ:	ɓɨ:n
f				*	*	fien	fɨ:	
v		vit	ve:	*	*	vieŋ	*	*
t		ti:p		*	*	tiem	tɨ:	tɨ:ŋ
d	di:	dik	de:	*	*	diek	dɨ:	dɨ:ŋ
n	ni:	nip		*	*	niep	nɨ:	
θ	θi:	θim	θe:	*	*	θien	θɨ:	θɨ:ŋ
ɗ		ɗip	ɗe:	*	*	ɗieŋ	ɗɨ:	ɗɨ:t
l/ð		ðit		*	*	ðieŋ	ðɨ:	ðɨ:ŋ
dz	dzi:	*	dze:	*	*	*	dzɨ:	dzɨ:ŋ
s	si:	*	se:	*	*	*	*	*
j	*	jit	*	jeŋ	*	*	*	*
x	*	xik	*	xet	*	*	*	*
y	yi:	yit	ye:	yek	*	*	yɨ:	yɨ:ŋ
ɲ	*	*	*	*	*	*	*	*
k	ki:	kin	ke:	*	*	kien	kɨ:	
g	gi:	gim	ge:	*	*	giew	gɨ:	
ŋ		ŋi:ŋ		*	*	ŋien	*	*
h	hi:	hiŋ	he:	*	*	hiet	hɨ:	hɨ:k
p^y	*	*	*	*	*	*		
b^y	*	*	*	*	*	*		
$ɓ^y$	*	*	*	*	*	*		
m^y	*	*	*	*	*	*	*	*
g^y	*	*	*	*	*	*	*	*
h^y	*	*	*	*	*	*	*	*
k^w	k^wi:	*	k^we:	*	*	*	*	*
g^w		*	g^we:	*	*	*	*	*
$ŋ^w$		*	$ŋ^w$e:	*	*	*	*	*
$θ^w$		*	$θ^w$e:	*	*	*	*	*
$ɗ^w$		*	$ɗ^w$e:	*	*	*	*	*
l^w		*	l^we:	*	*	*	*	*
s^w		*	s^we:	*	*	*	*	*

From the above lists, it is not difficult to see that the distribution of the phonological elements and expressions within a single, non-analytic domain in HT is not free but strictly constrained. Several observations can be made from the above tables. First, when there is only the element A in N_1, as in /a/ and /ə/, it can license almost all the onsets save /j/ and /x/. Secondly, as has been stated in the previous chapter, headless expressions cannot spread in HT. Thus in the above tables, the sequence of O_1 followed by /ə/ when O_2 is empty is ruled out, leaving the column of /ə/ blank. Thirdly, the above tables show the cases when O_1 is empty, e.g. [in], [op], etc. It is obvious that when O_1 and O_2 cannot both be p-licensed. The fourth observation from (4) is that when N_1 contains the element I or U, as in /i/, /e/, /u/ and /o/, the distribution is more restricted. This suggests that the elements I and U, unlike A, must have some constraints in licensing onsets. In the following subsections, I will provide more data with the elements I and U, so that respective constraints and onset representations can be established. Now, let us begin by discussing the constraint with the element I in HT.

4.2.2.1 I-based onsets

Before proposing the I-constraint of HT, let us have a look at the distribution of onsets in O_1 with the nuclei containing element I, i.e. /i/, /e/ and /ie/. The

distribution will provide more evidence to the discussion of /ie/ in Chapter Three and further confirms that it is in fact /^{y}e/, which is derived from /e/. The following table is also divided into two cases—when O_2 is empty and when it is filled.

(5) Distribution of the element I

	i:	i (:) + O_2	e:	e + O_2	ie:	ie + O_2
	*	in	*	eŋ	*	*
p	pi:	pi:k		*	*	pieŋ
b	bi:	bit	be:	*	*	biek
m	mi:	mi:ŋ	me:	*	*	mien
ɓ		ɓi:n	ɓe:	*	*	ɓiet
f				*	*	fien
v		vit	ve:	*	*	vieŋ
t		ti:p		*	*	tiem
d	di:	dik	de:	*	*	diek
n	ni:	nip		*	*	niep
θ	θi:	θim	θe:	*	*	θien
ɗ		ɗip	ɗe:	*	*	ɗieŋ
ð		ðit		*	*	ðieŋ
l	li:	lim	le:	*	*	lieŋ
dz	dzi:	*	dze:	*	*	*
s	si:	*	se:	*	*	*
j	*	jit	*	jeŋ	*	*
x	*	xik	*	xet	*	*
y	yi:	yit	ye:	yek	*	*
ɲ	*	*	*	*	*	*
k	ki:	kin	ke:	*	*	kien
g	gi:	gim	ge:	*	*	giew
ŋ		ŋi:ŋ		*	*	ŋien
h	hi:	hiŋ	he:	*	*	hiet
p^y	*	*	*	*	*	*
b^y	*	*	*	*	*	*
$ɓ^y$	*	*	*	*	*	*
m^y	*	*	*	*	*	*
g^y	*	*	*	*	*	*
h^y	*	*	*	*	*	*
k^w	k^wi:	*	k^we:	*	*	*
$ŋ^w$		*	g^we:	*	*	*
g^w		*	$ŋ^w$e:	*	*	*
$θ^w$		*	$θ^w$e:	*	*	*
$ɗ^w$		*	$ɗ^w$e:	*	*	*
l^w		*	l^we:	*	*	*
s^w		*	s^we:	*	*	*

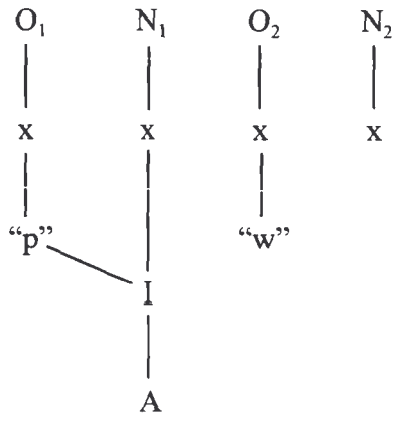
There are two obvious facts in the above table. First, the distribution of /e/ and /ie/ alternates with each other, depending on whether O_2 is filled. Secondly, there are several exceptions to this alternation. The answer to the first fact is the I-Spreading Constraint I proposed in Chapter Three, which is repeated below for the reader's convenience.

(6) I-Spreading Constraint in HT

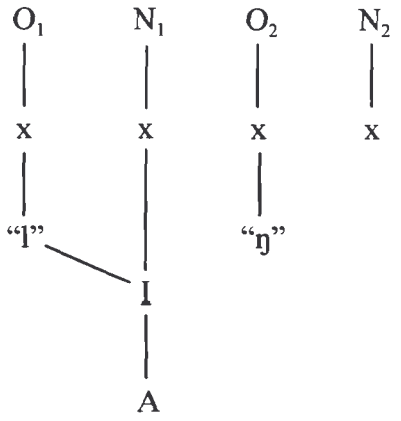
The element I in N_1 must spread. If it cannot spread to the right, it spreads to the left.

This constraint helps us to realize that the correct forms for /piew/, / lieŋ/, etc.—words that contain /ie/, are in fact /pʸew/ and /lʸeŋ/, with /y/ derived from N_1. This explains why /ie/ has such a restricted distribution—it is derived from the lexical nucleus /e/ when O_2 is filled. Examples are provided below to illustrate the spreading process.

(7) a. [pʸew] 'to throw away'



b. [lʸeŋ] 'afternoon meal'



The above structures demonstrated how the I-spreading process takes place.

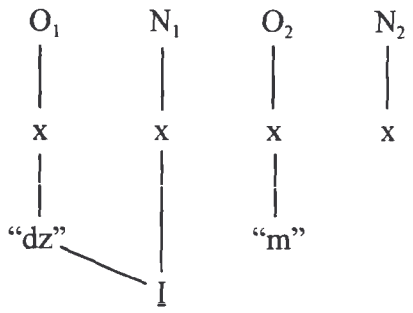
Now we need to explain the exceptions. I propose that the onsets that cannot be

followed by /ie/ are those that contain the element I themselves. The I-Spreading Constraint cannot apply here because these expressions themselves already have the element I. It makes no sense for them to receive another I from other positions. It would also violate the OCP (Obligatory Contour Principle)[4] if a given element appears more than once in a phonological expression. Thus, we can conclude that onsets /y/, /ɲ/, /j/, /x/, /pʸ/, /bʸ/, /mʸ/, /ɓʸ/, /gʸ/ and /hʸ/ contain the element I, but not onsets /dz/ and /s/.

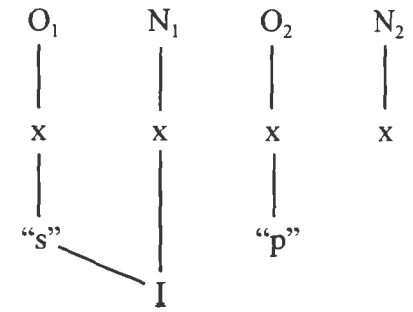
The reason that /dz/ and /s/ are excluded is that it is not just /ie/ that cannot follow them. The table of the complete distribution provided in (4) tells us that these two onsets cannot be followed by any nucleus that contains the element I when O_2 is filled. On the other hand, we can find that the distribution of onsets /j/ and /x/ is just the contrary. They can only appear before nuclei that contain I when O_2 is filled. The fact that the distribution of these two pairs of onsets is complementary leads us naturally to another fact, i.e. they do not form any phonological contrasts or minimal pairs. This behavior implies that /j/ and /x/ are derived onsets with the element I spreading to /dz/ and /s/. The spreading is shown in the following configurations.

[4] See Leben (1973) and Charette (1990) for the details.

(8) a. [jim] 'taste'



b. [xip] 'to smell'

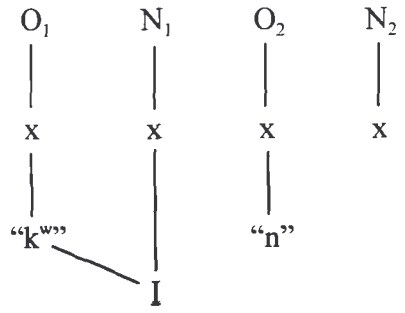


The above two pairs of examples demonstrated how /j/ and /x/ are derived from /dz/ and /s/ due to I-spreading from N_1.

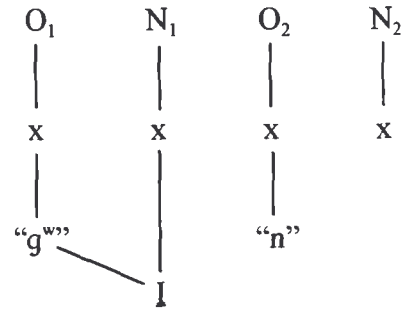
The I-Spreading Constraint not only explains the complementary distribution of /e/ and /ie/ and the derivation of /j/ and /x/. It is also the answer to why /k^w/, /g^w/, /$ŋ^w$/, /$θ^w$/, /d^w/, /s^w/ and /l^w/ cannot be followed by an N_1 containing I when O_2 is filled. In (5), we can find that these onsets are followed by /i/ and /e/ only when O_2 is empty.

This is due to the application of the I-Spreading Constraint when O_2 is filled. Thus, instead of having formations like *[k^{w}en], *[θ^{w}in], etc., we find [küeŋ] 'circle', [güin] 'fist', [lüen] 'roll', etc. The derivation follows.

(9) a. [küeŋ] 'mussel'



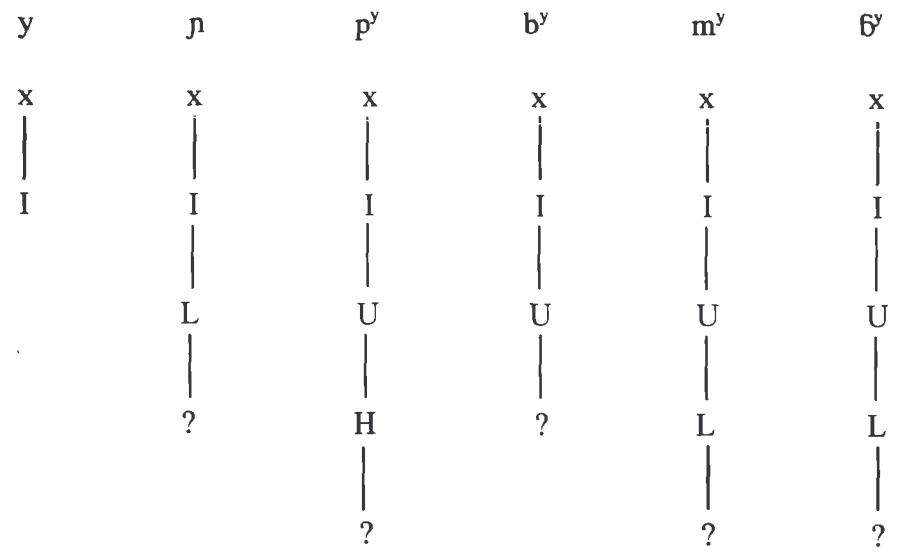
b. [güin] 'fist'

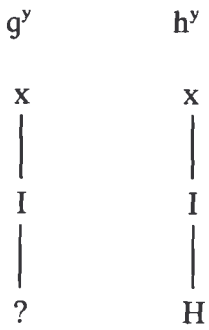


Now we can conclude that /y/, /ɲ/, /p^{y}/, /b^{y}/, /m^{y}/, /ɓ^{y}/, /g^{y}/ and /h^{y}/ are lexical onsets that contain the element I and /j/ and /x/ are derived onsets due to I-spreading. Thus,

leaving the discussion of headship to 4.2.2.3, we can set up the preliminary representations for the I-based onsets.[5]

(10)





Having defined onsets that contain I, now let us turn to the onsets that contain U.

[5] See the following subsections for the representations of /p/, /b/, /m/, /ɓ/, /g/ and /h/.

4.2.2.2 U-based onsets

Before discussing the onsets that contain the element U, let us look at the distribution of /u/ and /o/, the nuclei that contain U. They are also divided into two groups—when O_2 is empty and when O_2 is filled. The symbols in the table carry the same meaning as in (4) and (5).

(11) Distribution of the element U

	u:	u (:) + O_2	o:	o (:) + O_2
	*	um	*	op
p		pu:k	po:	poŋ
b	bu:	buy	bo:	bop
m	mu:	mum	mo:	mo:n
ɓ		ɓuŋ	ɓo:	ɓo:t
f		fuk		fow
v	*	*	*	*
t	tu:	tu:m	to:	tom
d	du:	duk	do:	dop
n	nu:	nu:k	no:	no:k
θ		θu:n		θo:n
ɗ		ɗuŋ	ɗo:	ɗot
ð	ðu:	ðu:ŋ		ðot
l	lu:	lu:ŋ	lo:	lok
dz	dzu:	dzuk		dzop
s	su:	suk	so:	so:n
y	yu:	yuk	yo:	yok
ɲ		ɲup		ɲo:m
k		kum	ko:	kom
g	gu:	guk	go:	go:t
ŋ	ŋu:	ŋu:m	ŋo:	ŋo:k
h	hu:	hun	ho:	hop
p^y		$p^yu:m$		p^yoy
b^y	$b^yu:$	$b^yu:y$		b^yok
$ɓ^y$	$ɓ^yu:$			$ɓ^yo:k$
m^y	*	*	*	*
g^y	$g^yu:$			
h^y	$h^yu:$			
k^w	*	*	*	*
$ŋ^w$	*	*	*	*
g^w	*	*	*	*
$θ^w$	*	*	*	*
$ɗ^w$	*	*	*	*
l^w	*	*	*	*
s^w	*	*	*	*

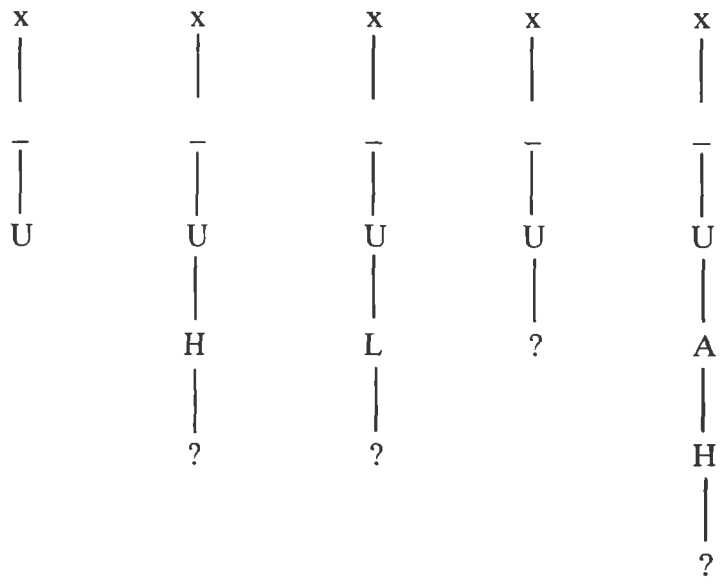
To describe the distributional constraints on U, I assume that the following constraint holds at the L-structure of HT.

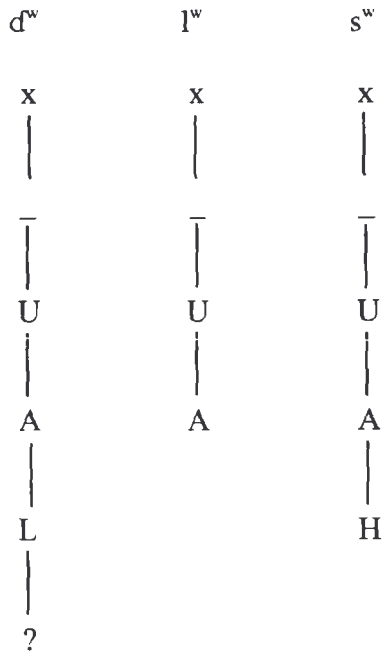
(12) The U-Constraint

If O_1 and N_1 both contain U, then O_1 must be headed.

Given the above constraint and the restricted distribution in (11), we can easily conclude that onsets /v/, /k^w/, /$ŋ^w$/, /g^w/, /$θ^w$/, /d^w/, /l^w/ and /s^w/ contain U and are headless, as they cannot precede the nuclei containing U in any case. The representations are thus proposed as follows.

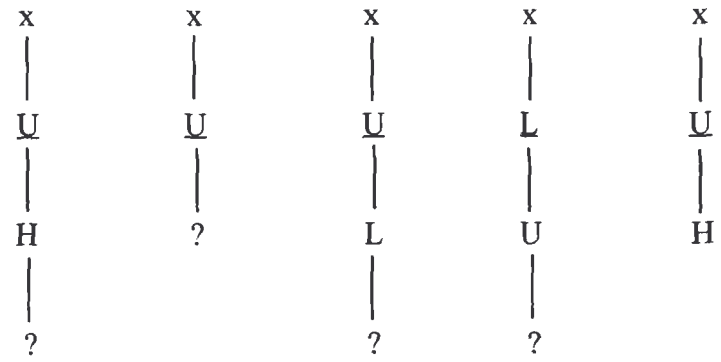
(13)





As for the other onsets in the same set with /v/, namely /p/, /b/, /m/ /ɓ/ and /f/, I propose that they also have U[6]. As they can be followed by N_1 containing U, according to (12), they must be headed. Their representations are thus proposed below.

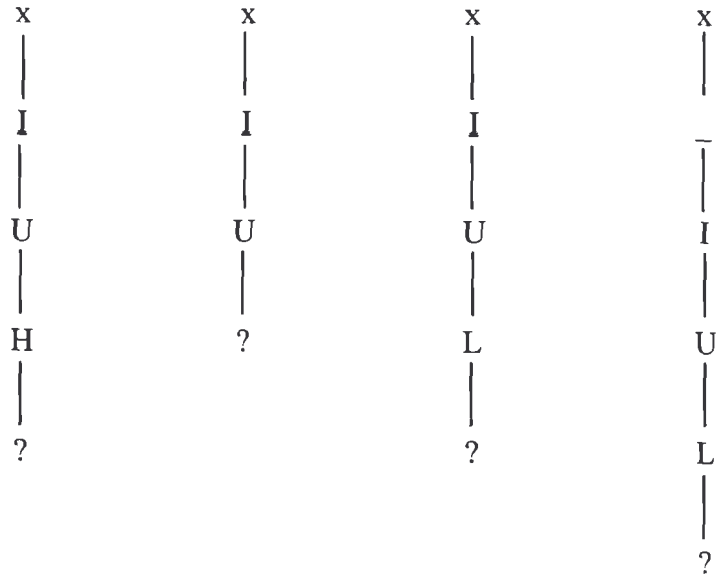
(14)



[6] In the next chapter, discussion of the spreading from O_2 to N_1 will provide further evidence to prove that these onsets contain the element that spreads to the head of N_1.

In addition to the above onsets, onsets /pʸ/, /bʸ/, /mʸ/ and /ɓʸ/ mentioned in the previous subsection also contain U. Their distribution in (11) shows that /pʸ/, /bʸ/ and /ɓʸ/ can be followed by N_1 containing U while /mʸ/ cannot. Following this and the constraint proposed in (12), I suggest that /pʸ/, /bʸ/ and /ɓʸ/ are headed and /mʸ/ headless. The representations are proposed below.

(15)



4.2.2.3 The empty nucleus constraint

A glance at the distribution when N_1 is empty shows that it can license only a restricted set of onsets both in O_1 and in O_2. I will only provide its distribution with

O_1 in this chapter and leave its distribution with O_2 to the next, which focuses on the interaction between onset and nucleus. The symbols in the table carry the same meaning as in (4) and (5).

(16) Distribution of the empty nucleus

	ɨː	ɨː + O_2
p	pɨː	
b	bɨː	bɨːŋ
m	mɨː	mɨːŋ
ɓ	ɓɨː	ɓɨːn
f	fɨː	
v	*	*
t	tɨː	tɨːŋ
d	dɨː	dɨːŋ
n	nɨː	
θ	θɨː	θɨːŋ
ɗ	ɗɨː	ɗɨːt
ð	ðɨː	ðɨːŋ
l	lɨː	lɨːk
dz	dzɨː	dzɨːŋ
s	*	*
y	yɨː	yɨːŋ
ɲ	*	*
k	kɨː	
g	gɨː	
ŋ	*	*
h	hɨː	hɨːk
p^y		
b^y		
$ɓ^y$		
m^y	*	*
g^y	*	*
h^y	*	*
k^w	*	*
g^w	*	*
$ŋ^w$	*	*
$θ^w$	*	*
$ɗ^w$	*	*
l^w	*	*
s^w	*	*

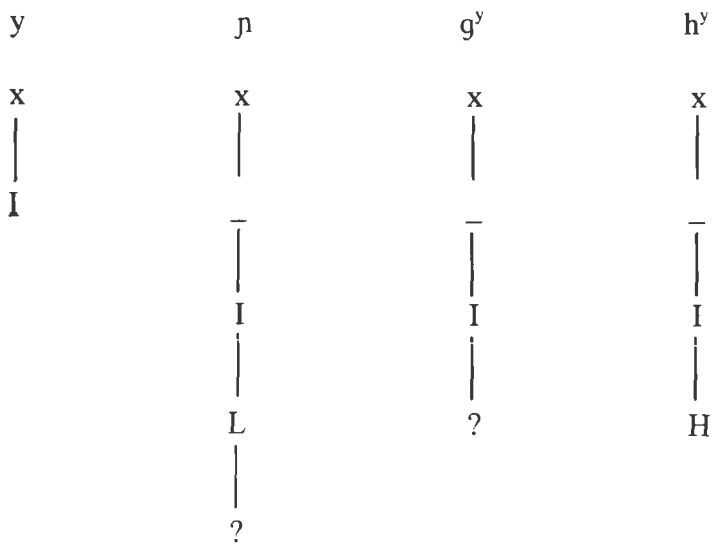
In the above table, it is not difficult to see that when N_1 is empty, it has a restricted distribution with O_1 that contain U or I. Onsets /v/, /k^w/, /g^w/, /ŋ^w/, /θ^w/, /ɖ^w/, /l^w/ and /s^w/, which are proposed to be headless in the previous subsection, are ruled out before the empty nucleus. Similarly, of the set of onsets containing I, namely, {y, ɲ, p^y, b^y, m^y, ɓ^y, g^y, h^y}, only /y/ can appear before the empty N_1. To capture these distributional gaps as well as to define the onsets that can be licensed by an empty nucleus, I propose the following constraint.

(17) The Empty Nucleus Constraint

When N_1 is empty, If O_1 contains A, I or U, it must be headed. Otherwise it is headless.

This constraint further proves that onsets /p/, /b/, /m/, /ɓ/ and /f/ are headed, as they can be followed by the empty nucleus, while /v/, /k^w/, /g^w/, /ŋ^w/, /θ^w/, /ɖ^w/, /l^w/ and /s^w/ are headless. In addition, it helps us to conclude that, of the remaining I-based onsets we have examined above, /y/ is headed while /ɲ/, /g^y/ and /h^y/ are headless. Thus, /y/ can be followed by the empty nucleus and /ɲ/, /g^y/ and /h^y/ cannot. Their representations are proposed as follows.

(18)



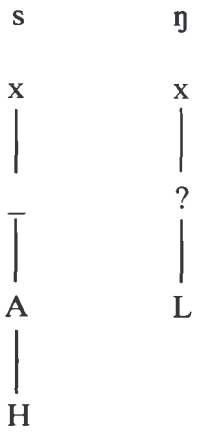
Now we are left with the remaining two sets of onsets, {t, d, l, n, ɗ, θ, ð, dz, s} and {k, g, ŋ, h}. From the discussion of the I- and U-based onsets in the previous subsections, we know that these two sets do not contain I, nor U. Yet they must be differentiated by the presence of an element, which by a process of elimination can only be A. Based on the previous works of GP[7], I assume that the set {t, d, l, n, ɗ, θ, ð, dz, s} contains A and the set {k, g, ŋ, h} does not.

Next we need to determine the headedness of these two sets of onsets. The distribution table (16) shows that, of these two sets, /s/ and /ŋ/ cannot be followed by the empty nucleus. The above discussion suggests that /s/ contains A and /ŋ/ does not. Constraint (17) proposes that if O_1 contains A, I or U, it must be headed in order to be followed by an empty nucleus. Its contraspositive (namely, if O_1 containing A, I or U cannot be followed by the empty nucleus, it must be headless)

[7] For example, Goh (1996a, b), Kim (1996) and Pan (2000), just to name a few.

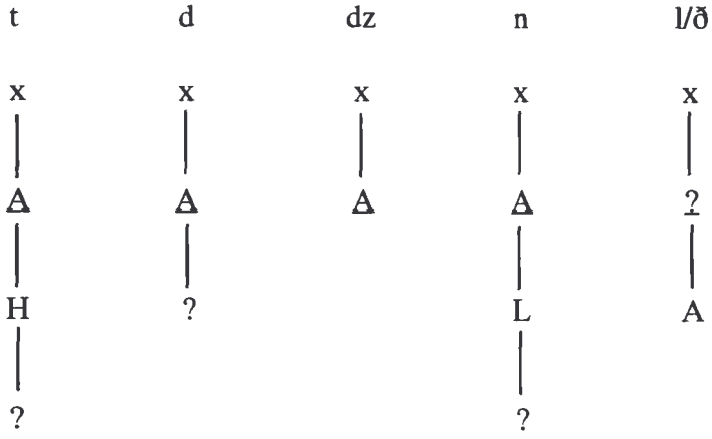
leads us to conclude that /s/ is headless. On the other hand, constraint (17) also proposes that when O_1 does not contain A, I or U, it must be headless in order to be licensed by the empty nucleus. It then follows that /ŋ/ is headed, as it cannot appear when N_1 is empty. The representations for /s/ and /ŋ/ are suggested below.

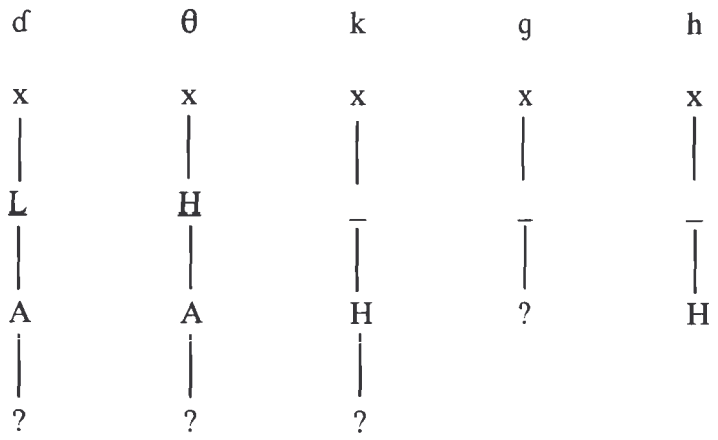
(19)



Finally I propose the representations of the remaining onsets as follows.

(20)





This concludes the onset system and the phonological representations for all the onsets in HT. There are still some other distributional constraints between the onset and the nucleus and between the two onsets in HT. As this chapter is primarily about the onset system of HT and the phonological representations for the onsets, I will present a specific discussion on the other constraints in the next chapter.

4.3 Summary

In this chapter, I have presented a detailed discussion on the onset system of HT and the phonological representations for the lexical onsets. It has been shown that some of the surface onsets in HT are derived. This chapter has also proposed the following constraints.

1. Constraints in the onset system of HT

i. The U-Constraint

If O_1 and N_1 both contain U, then O_1 must be headed.

ii. The Empty Nucleus Constraint

When N_1 is empty, If O_1 contains A, I or U, it must be headed. Otherwise it is headless.

Based on the above constraints, the phonological representations for the lexical onsets in HT are discovered to be:

2. Phonological representations for the lexical onsets in HT

i.	p ({H, ?}, U)	b ({?}, U)	m ({L, ?}, U)
	ɓ ({U, ? }, L)	f ({H}, U)	v/w ({U}, _)
ii.	t ({H, ?}, A)	d ({?}, A)	n ({L, ?}, A)
	ɗ ({A, ?}, L)	θ ({A, ?}, H)	l/ð ({A}, ?)
	dz ({}, A)	s ({A, H}, _)	
iii.	y ({}, I)	ɲ ({I, L, ?}, _)	
iv.	k({H, ?}, _)	g ({?}, _)	ŋ ({L}, ?)
	h ({H}, _)		
v.	p^y ({U, H, ?}, I)	b^y ({U, ?}, I)	m^y ({I, U, L}, _)
	$ɓ^y$ ({U, L, ?}, I)	g^y ({I, ?}, _)	h^y ({H, I}, _)

vi. k^w ({U, H, ?}, _) g^w ({U, ?}, _) $ŋ^w$ ({U, L, ?}, _)

$θ^w$ ({U, A, H, ?}, _) d^w ({U, A, L, ?}, _) l^w ({A, ?}, _)

s^w ({U, A, H}, _)

As this is the first phonological study on HT, I am sure that there will be more discoveries with more research carried out. On the other hand, the chapter has provided some of the constraints that may be used to illustrate the phonological representations of onsets in HT. More constraints on onsets and nuclei will be provided in the next chapter.

Chapter Five

Inter-Constituent Interactions

5.0 Introduction

Chapter Three and Chapter Four have discussed the nuclear and onset system of HT. In the process of the discussion, some constraints, which are based on the interactions between onset and nucleus, have been introduced. In order to present a comparatively clear picture of the phonological system of HT, I will introduce more constraints between constituents and explain how they interact with each other so as to give a systematic account on the inter-constituent interactions of HT.

This chapter is divided into four sections. Interaction between the two nuclei in a phonological domain of HT will be introduced in the first section. Then the following section, 5.2, will focus on the interactions between onset and nucleus. It is further divided into three parts. The first part will discuss the constraint between O_2 and N_2. The second part summarizes the interactions between O_1 and N_1, as it has been demonstrated in detail in the last two chapters. Section 5.2.3, will

introduce the interactions between N_1 and O_2. Although discussions of the nuclear and onset system in the previous chapters did not mention the interactions of N_1 and O_2, data in HT show an equally active interaction between these two constituents which accords with the whole phonological system of the language very well. Both 5.2.2 and 5.2.3 consist of a separate discussion of spreading and constraints. The discussion of spreading concerns how elements spread, which also includes the direction of spreading. Discussions of constraints, on the other hand, will show how the components of constituents influence each other and will provide explanations for the restricted distribution. Section 5.3 discusses the interaction between the two onsets, O_1 and O_2. Finally, section 5.4 will give a summary of the discussions provided in this chapter.

5.1 Interaction Between Nuclei

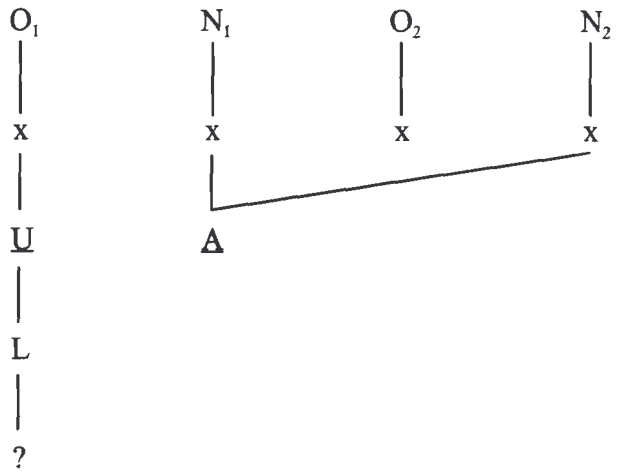
In Chapter Two, I have claimed that HT has a basic phonological template of two pairs of onsets and nuclei. It has also been shown that N_1 is the head of the template, from which N_2 inherits the licensing power so as to license O_2. Data in HT show that N_2 cannot have lexical material, i.e. it is lexically empty, but it may receive material spread from N_1. There is one requirement for the spreading—when O_2 is p-licensed. Thus the interaction between N_1 and N_2 can be proposed as follows.

(1) Interaction between nuclei in HT

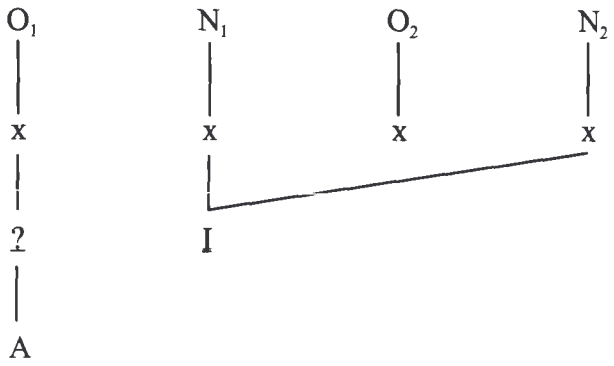
If N_1 is a headed expression, it must spread to N_2 when O_2 is p-licensed.

The above interaction defines the spreading from N_1 to N_2 as obligatory when O_2 is p-licensed. In addition, it requires that N_1 must be a headed expression in order to spread. There are two lexical nuclei in HT that do not have head in their phonological representations—/ə/ and /ɨ/. /ə/ is never found to be long, which confirms the above proposal. The reason that /ɨ/ is always found to take two positions when O_2 is not filled in HT has been provided in Chapter two and Chapter three. It has been shown that it is not due to spreading, but rather that, to show their presence, one of the pair of onset and nucleus must not be p-licensed and thus be audible. Examples with all the lexical nuclei in HT are provided below to demonstrate the rule proposed in (1).

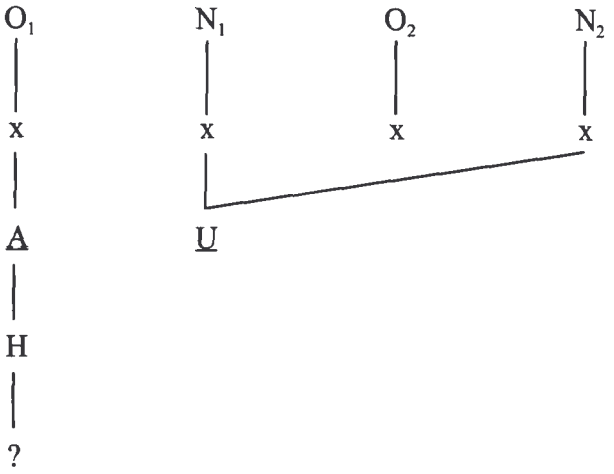
(2) a. [ma:] 'horse'



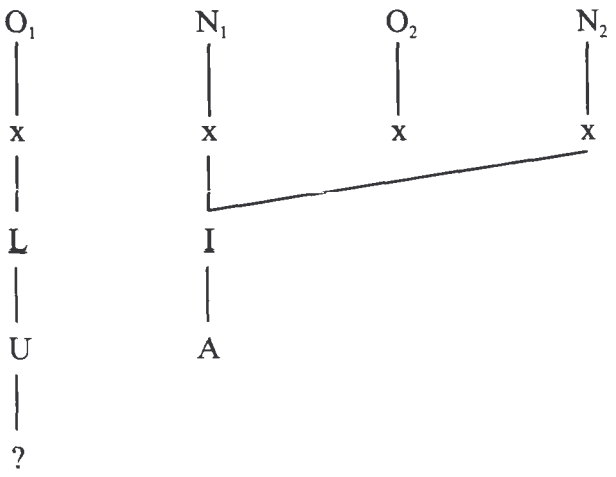
b. [li:] 'to aim'



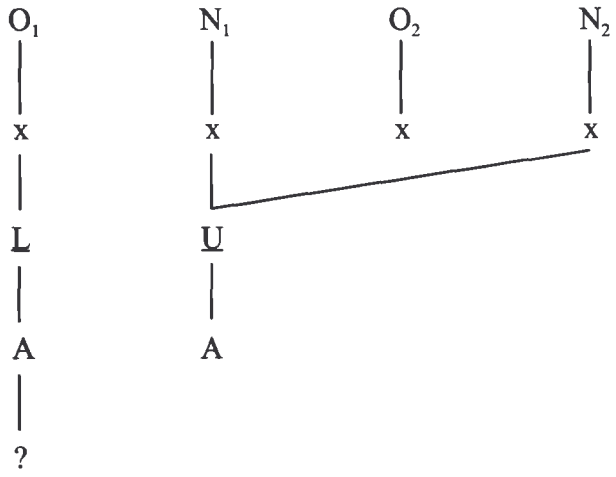
c. [tu:] 'head'



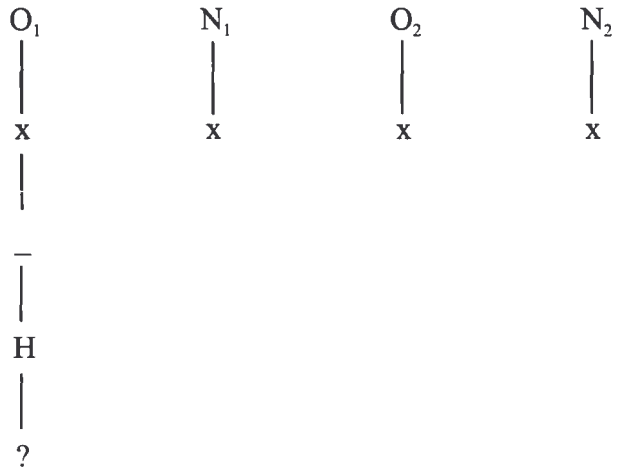
d. [ɓe:] 'sheep'



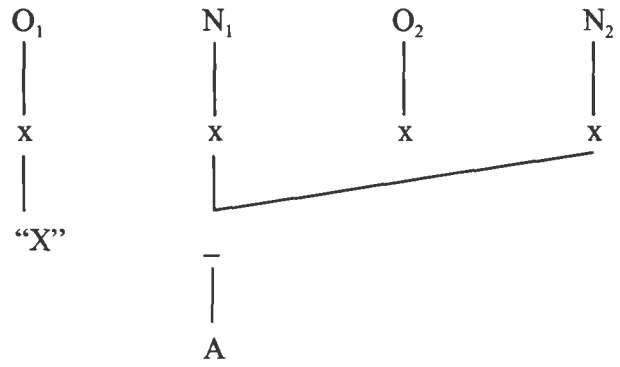
e. [ɗo:] 'to hide/avoid'



f. [kɬ:] 'eggplant'



g. *[Xə:]



Examples (2a) to (2e) showed how the process of headed N_1 spreading to N_2 takes place when O_2 is p-licensed. In these five examples, all the first nuclei are taken by a headed expression, either one of the simplex expressions—({}, A), ({}, I)

and ({}, U) as in (2a) to (2c), or a complex expression—({A}, I) and ({A}, U) in (2d) and (2e). On the other hand, the second onsets in these five examples are all p-licensed. Thus the two conditions for the interaction between N_1 and N_2 are met. In this case, N_1 spreads to N_2.

In (2f), however, there is no spreading taking place. N_1 is an empty position. As it is the head of the domain that initiates licensing, it must not be p-licensed. On the other hand, because it is empty, it has nothing to spread. Thus N_2 remains empty. As O_2 is also empty, there are three empty positions in this phonological domain. As discussed above, a pair of onset and nucleus cannot both be p-licensed, otherwise there would be no way to know their presence. Thus, in (2f), either O_2 or N_2 must not be p-licensed. Following its performance in the augmented template, I propose that N_2 is not p-licensed, which results in that two empty nuclei are heard in (2f).

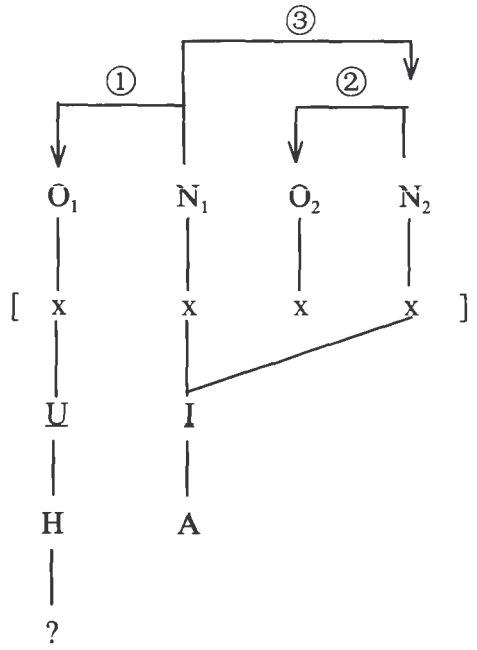
In (2g), "X" represents any given onset in HT, while N_1 is an empty-headed expression. As headless expressions cannot spread in HT, structures like (2g) are ruled out in this language. Now let us turn to next section and see how onset interacts with nucleus.

5.2 Interactions Between Onset and Nucleus

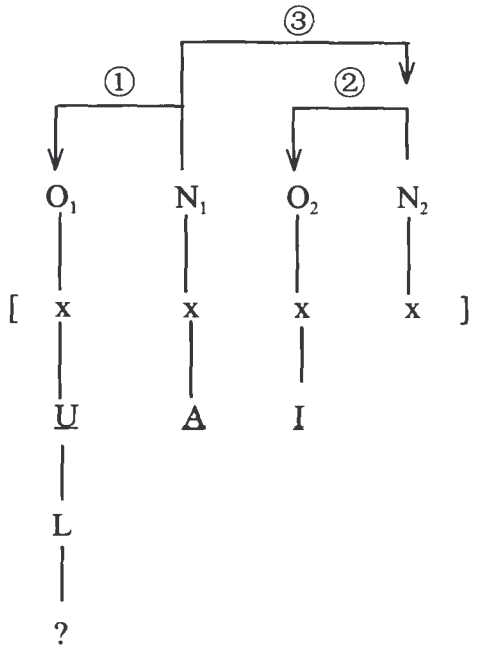
Besides the interaction between the nuclei introduced in the last section, HT shows a very active interaction between onset and nucleus. Unlike the interaction between nuclei, which consists of only spreading, the interaction between onset and nucleus include more constraints. There are also several types of spreading between onset and nucleus, but it is different from the one between nuclei. Spreading between the nuclei, as we have seen, includes the complete phonological expression of the constituent. Spreading between onset and nucleus, however, will concern only one element of a phonological representation.

Of the two sets of onsets and nuclei, there is no interaction between O_1 and N_2, as there is neither governing nor licensing relationship between them and they are not adjacent at any level. Examples are provided below to demonstrate the structure.

(3) a. [pe:] 'to smack'



b. [may] 'tree'



The above examples showed the two possible occasions with N_2. In (3a), N_2 is filled with the material spread from N_1, while in (3b), no spreading happens between the two nuclei and N_2 is p-licensed. Although the N_2 in the above two examples appears to have different materials at the surface level, the lexical structure remains the same. The numbers clearly indicate the relationship between constituents. Step ① shows the licensing relationship between N_1 and O_1. The licensing relationship between the second pair of onset and nucleus is indicated by ②. N_1 governs N_2 at the nuclear projection level as shown in ③. Besides these three relationships, the template shows no relationship between O_1 and N_2.

In the following, I will present three types of onset-nucleus interactions found in HT: O_2-N_2, O_1-N_1 and O_2-N_1. Now let us look at the interaction between O_2 and N_2 first.

5.2.1 Interaction between O_2 and N_2

Chapter Two and Three have demonstrated that O_2 and N_2 cannot both be interpreted in a phonological domain. The phonological template of HT requires that N_2 be lexically empty. The interaction constraint proposed in (1) implies that the spreading from N_1 to N_2 is dependent on the status of O_2. N_2 can receive the

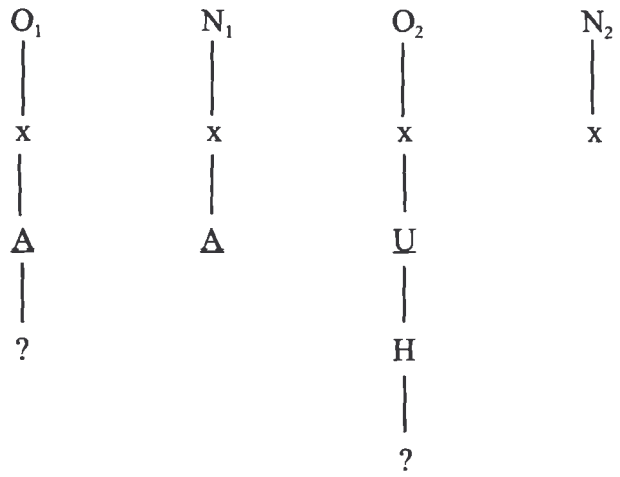
spreading from N_1 only when O_2 is p-licensed. On the other hand, when O_2 is filled, it blocks the spreading and N_2 has to remain lexically empty. Therefore from (1), we naturally derive the following constraint between O_2 and N_2.

(4) Constraint Between O_2 and N_2

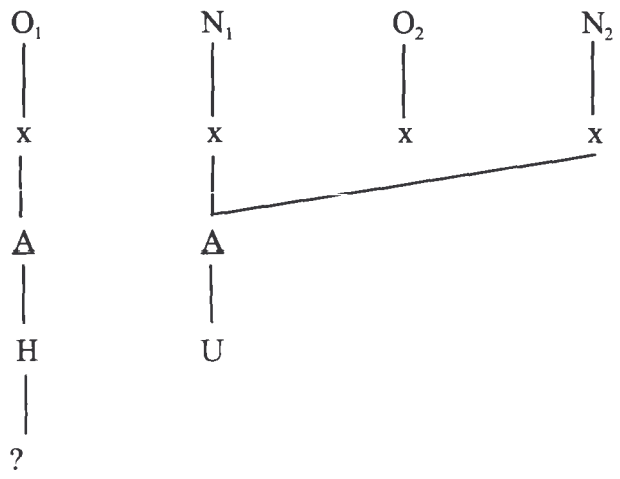
O_2 and N_2 cannot both be filled.

As one of them must be p-licensed in the phonological template of HT, no spreading can take place between O_2 and N_2. Examples are provided below to demonstrate the constraint between O_2 and N_2 proposed in (4).

(5) a. [da^p] 'to pile'



b. [to:] 'to sweep'



In (5a), O_2 is filled, which blocks the spreading from N_1 to N_2. N_2 thus remains lexically empty and is p-licensed. In (5b), N_1 spreads to N_2 because O_2 is p-licensed and is not in the way of spreading. In neither of the cases can O_2 and N_2 both be filled. Next, let us turn to see the interaction between N_1 and the two onsets.

5.2.2 Interaction between O_1 and N_1

In the last two chapters, the interaction between O_1 and N_1 has been introduced and has been used as a means of defining properties of phonological representations and discovering the nuclear and onset system of the language. The

following will provide a summary of the interaction between O_1 and N_1 based on the discussion in the last two chapters. First let us examine the spreading.

5.2.2.1 Spreading between O_1 and N_1

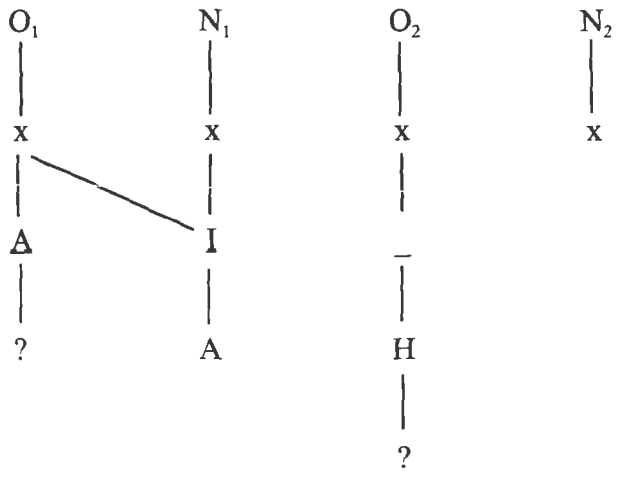
From the discussion in the last two chapters, it is not difficult to see that the spreading between the first pair of onset and nucleus is always one-way, i.e. from right to left. So O_1 does not spread to N_1, but the element I in N_1 may spread to O_1 when a condition is met, viz. when O_2 is filled. So we can propose the following rule of spreading between O_1 and N_1.

(6) Spreading between O_1 and N_1 in HT

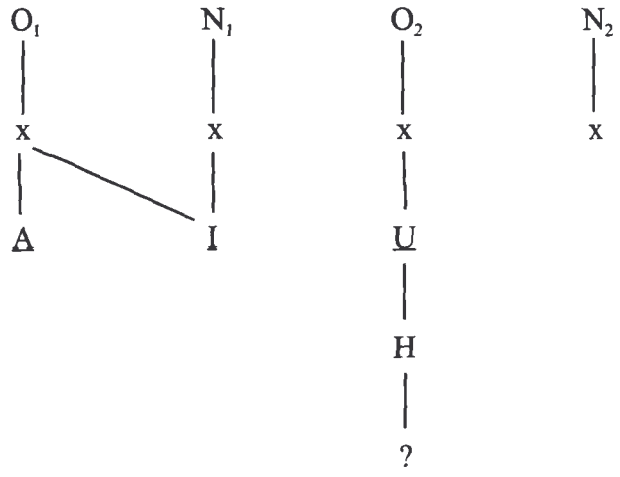
The element I in N_1 spreads to O_1 when O_2 is filled.

The following are examples illustrating how spreading proposed in (6) takes place.

(7) a. [$d^{y}e^{k}$] 'to put'



b. [ji^{p}] 'to ache / to hurt'



5.2.2.2 Constraints between O_1 and N_1

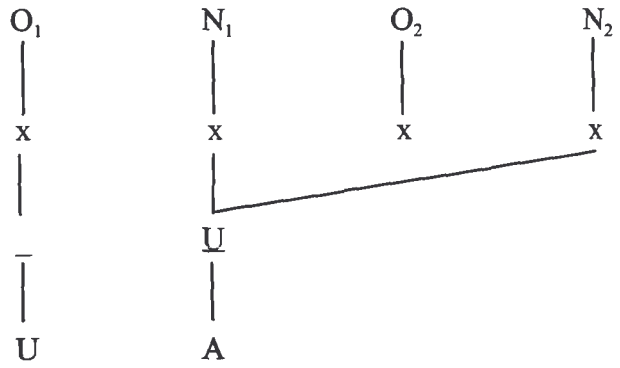
In Chapter Four, I proposed the U-Constraint in HT, which is repeated below for the reader's convenience.

(8) The U-Constraint in HT

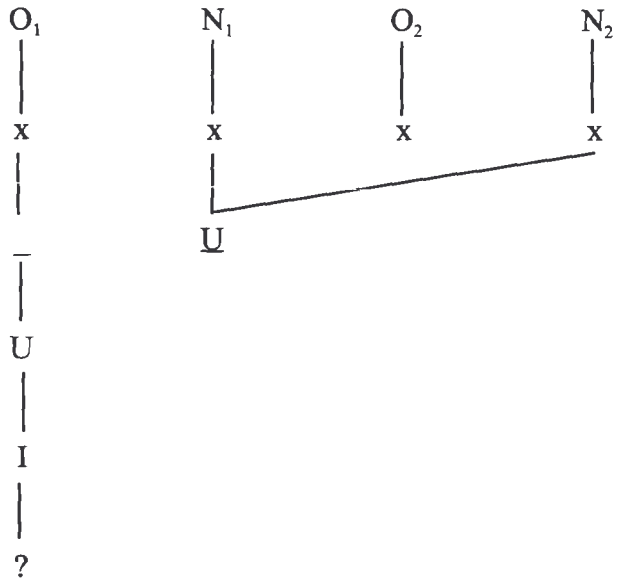
If O_1 and N_1 both contain U, then O_1 must be headed.

Given the above constraint, we will not expect to find O_1 to be headless when the two onsets both contain U. Examples are provided below to illustrate.

(9) a. * [vo:]



b. *[m^{y}u:]



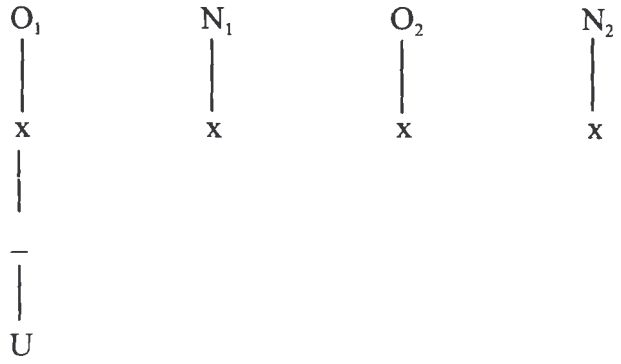
Chapter Four has also suggested the Empty Nucleus Constraint in HT repeated in the following.

(10) The Empty Nucleus Constraint

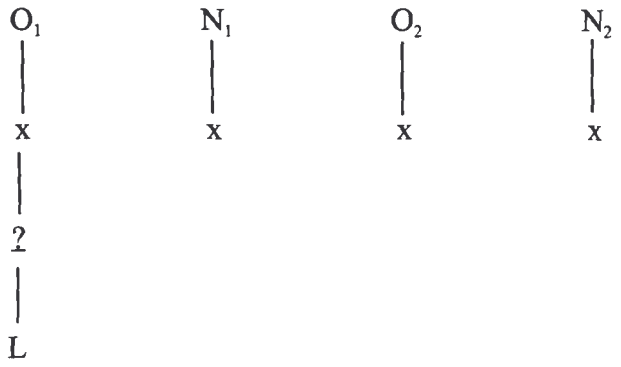
When N_1 is empty, If O_1 contains A, I or U, it must be headed. Otherwise it is headless.

The above constraint in fact divides onsets into two groups—those that contain A, I or U and those that do not. According to (10), in order to precede the empty N_1, onsets containing A, I or U must be headed, while the others must be headless. Examples are provided below to illustrate.

(11) a. *[vɬ:]



b. *[ŋɬ:]



5.2.3 Interactions between O_2 and N_1

The interactions between O_1 and N_1 have been used as a good way to define the onset and nuclear system of HT. The interactions between O_2 and N_1, to a certain

extent, are parallel to the interactions between the first pair of onset and nucleus. Typically, there are also leftward spreading of elements.

5.2.3.1 Spreading from O_2 to N_1

We have seen how I in N_1 spreads leftward to O_1 when O_2 is filled. From the investigation on the interactions between O_2 and N_1, I found that the left spreading of elements also takes place between these two constituents. It involves the elements U and I. Thus I propose the following process of spreading between O_2 and N_1 in HT.

(12) Spreading between O_2 and N_1 in HT

i. If U is the head of O_2, it must spread to the head of N_1; If U is in the operator position, the spreading is optional.

ii. I in O_2 may spread to the head of N_1.

The above proposal in fact implies that I and U in O_2 can spread only to a headless nuclear expression, as there can be at most one head in a phonological expression. A survey of the set of onsets that may appear in O_2, which are /w/, /y/, /ɣ /, /p/, /m/, /t/, /n/, /k/ and /ŋ/, shows that /w/, /p/ and /m/ contain U while /y/ contains I. Their representations are provided below again for the reader's convenience.

(13) w ({U}, _) p ({H, ?}, U) m ({L, ?}, U) y ({}, I)

According to (12), we will expect spreading to occur between a headless nuclear expression and these four onsets. As /p/ and /m/ are U-headed, their spreading to the head of N_1 should be obligatory; while the spreading from /w/ and /y/ to N_1 is optional according to (12). As we have discussed in Chapter Three, there are two headless nuclei in HT, /ə/ and /ɨ/. The next section that discusses the constraints between O_2 and N_1 will show that /ɨ/ cannot be followed by onsets containing I or U, so we are left with only nucleus /ə/. The data in HT consistently and clearly show a gap of sequences of /əm/ and /əp/, which is exactly what we expect to find. The distribution is as follows.

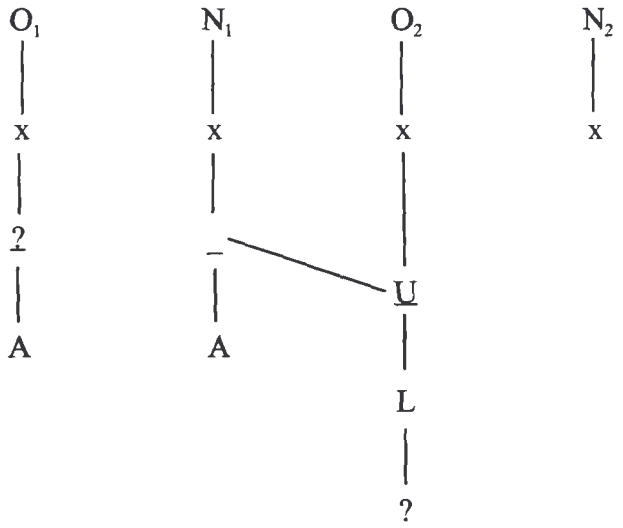
(14) Distribution of /ə/ with O_2

	y	w	p	m	n	t	k	ŋ	ɣ
ə	?	?	*	*	dən	lə^t	də́^{k}	tə̌ŋ	hə̌ɣ

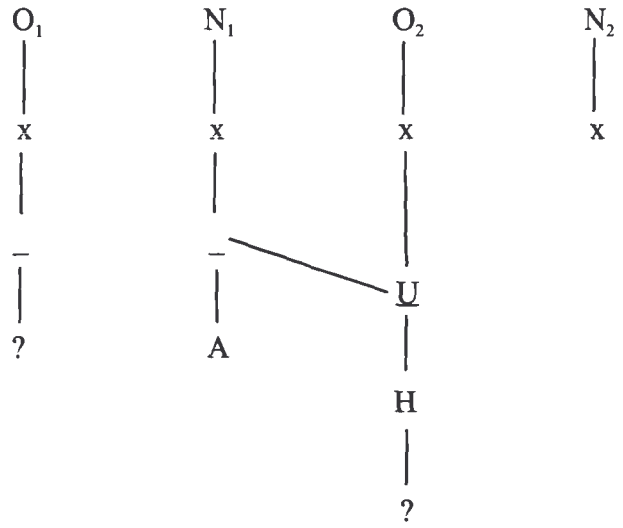
The situation with /əy/ and /əw/ is ambiguous, because, as suggested in (12), the spreading is optional. So I mark them with a question mark. There seems to be individual preferences among the informants of HT and they take both /əy/ and /ey/ and /əw/ and /ow/ as acceptable. Nevertheless, informants from Xintun village, which is next to Hongkui village, insist on the spread sequence, /ey/ and /ow/, and reject /əy/ and /əw/, where no spreading takes place. The following

examples demonstrate how the spreading proposed in (12) takes place and changes /əm/, /əw/, /əp/, and /əy/ to /om/, /ow/, /op/ and /ey/.

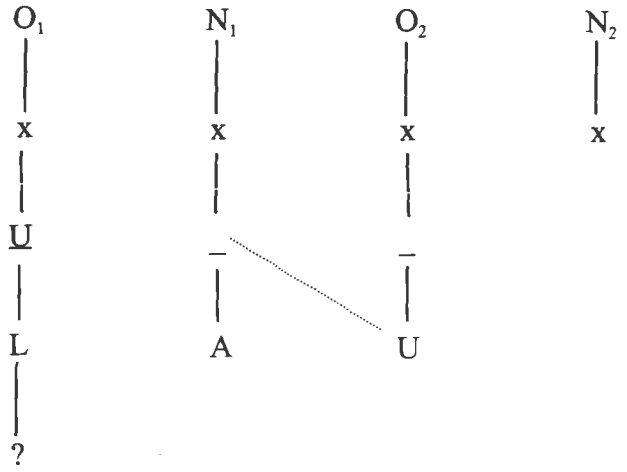
(15) a. [lom] 'to forget'



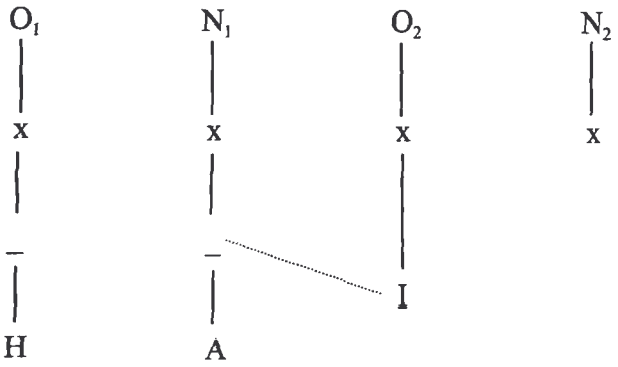
b. [go^{p}] 'frog'



c. [mow] 'pig'



d. [hey] 'to smell'



The spreading demonstrated in the above is parallel to the I-spreading found between O_1 and N_1, though the spreading from N_1 to O_1 is more conditional for it

requires O_2 to be filled. In (15c) and (15d), as the spreading is not obligatory, I used dotted line to indicate the optionality.

I mentioned above that although /ɨ/ does not have a head, there is no spreading of I and U to it. This is due to the constraint between O_2 and N_1. Now let us turn to the following subsection and see how the constraint between O_2 and N_1 takes place.

5.2.3.2 Constraints between O_2 and N_1

I proposed the spreading constraint between O_2 and N_1 in the above section. This explains the distributional gap of /ə/ with O_2 that contains U and I, which is shown in (14). In addition to /ə/, the empty nucleus /ɨ/ is found to have the same distributional constraint with O_2 that contains U or I. The distribution is shown below.

(16) Distribution of the empty nucleus with O_2

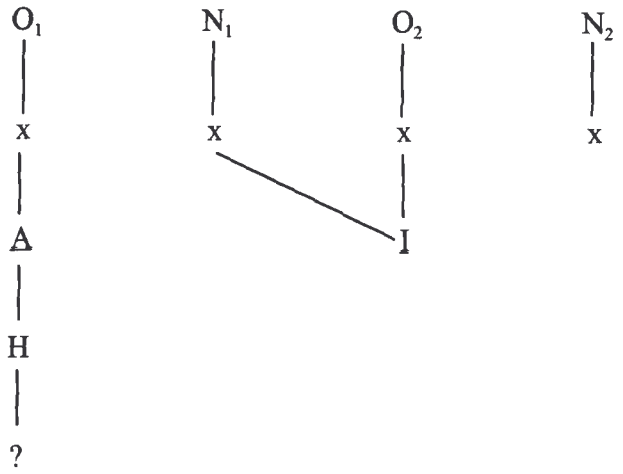
	y	w	p	m	n	t	k	ŋ	ɣ
ɨ	*	*	*	*	bɨ:n	dɨ:^{t}	dzɨ:^{k}	dzɨ:ŋ	*

This distribution table shows a similar gap to the one in (14), though the above distribution seems to be more restricted. There is no ambiguity as that has been

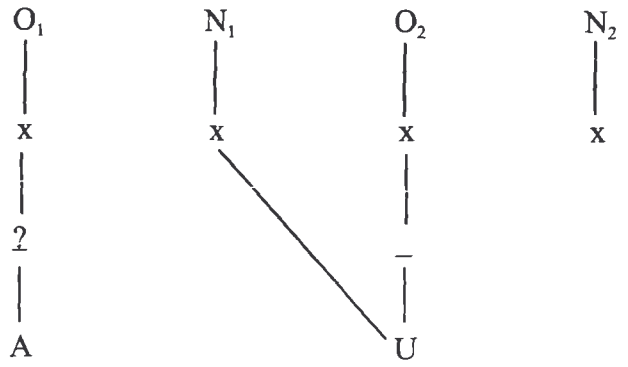
found with /əy/ and /əw/.

The reason that the empty N_1 cannot be followed by O_2 that contains I or U is in fact the spreading constraint proposed in (12). According to (12), I or U in O_2 spread to the head of N_1. Given the empty N_1 and O_2 that contains U or I, it would yield sequences as follows due to spreading.

(17) a. * [tiy]



b. * [luw]



It is not difficult to see that the above derivations will cause confusions to the whole phonological system of HT. For sequences like (17a) and (17b), it would be impossible to tell their differences with words like [ti:] and [lu:]. On the other hand, if we propose that the lexical structures for [ti:] and [lu:] were in fact as shown in (17), this would mean that the /i/ and /u/ in these cases are derived. Then we would need to explain why the nuclei /i/ and /u/ cannot follow O_1 in HT. Obviously no such explanation or constraint can be found in HT. So, to avoid such confusions, I propose that the following constraint between O_2 and N_1 holds at the L-structure of HT.

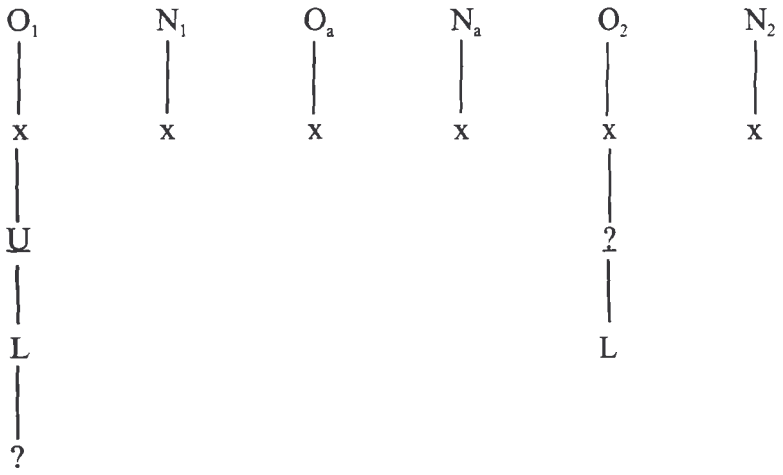
(18) Constraint between O_2 and N_1 in HT

When N_1 is an empty nucleus, O_2 cannot contain I or U.

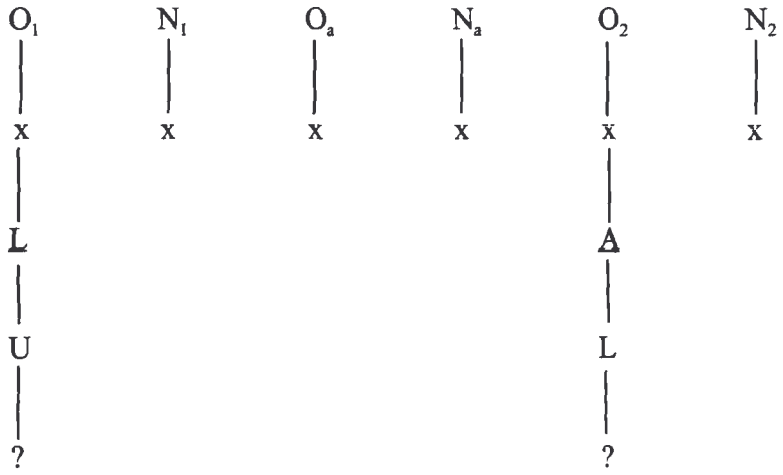
As the O_2 that contains I or U is not allowed to follow the empty N_1, the spreading

proposed in (12) will not be able to take place when N_1 is empty. Thus there will not be confusions like those mentioned above. More examples are provided below to illustrate the grammatical sequences of the empty N_1 with O_2.

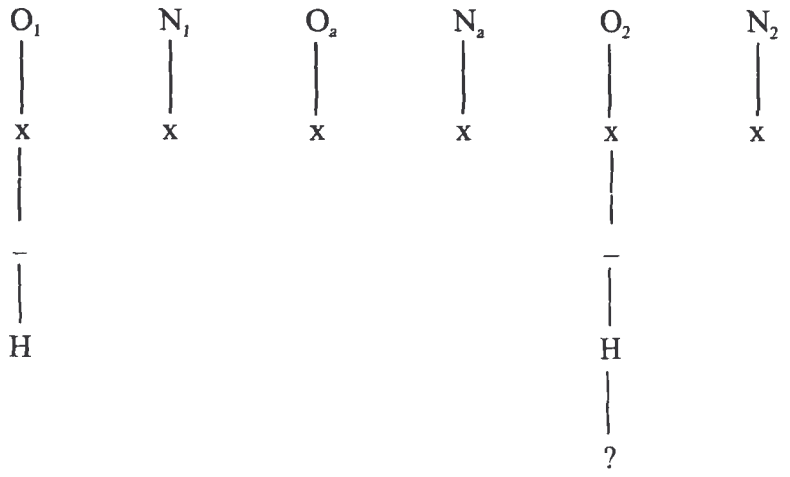
(19) a. [mɨ:ŋ] 'brook'



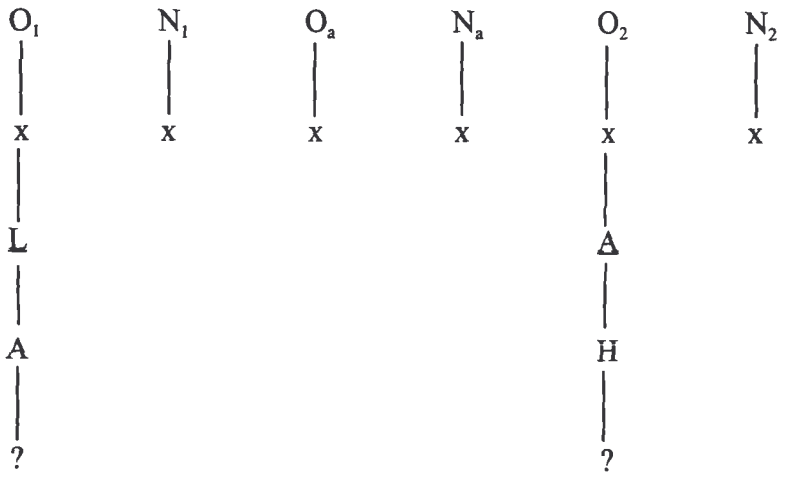
b. [bɨ:n] 'month'



c. [hɬː^{k}] 'gum'



c. [dɬː^{t}] 'hot'



5.3 Interaction Between Onsets

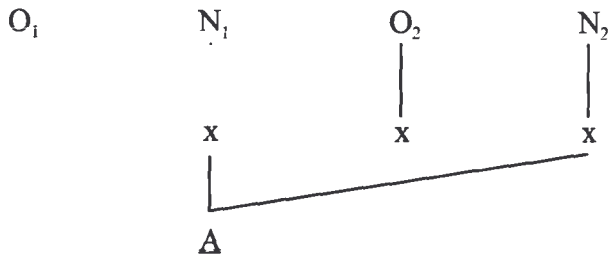
In the phonological template of HT, there is not much interaction between O_1 and O_2 due to the presence of N_1. I have discussed in Chapter Two that N_1 is the head of the domain and can never be p-licensed. Thus there cannot be spreading between the two onsets. However, in Chapter Two, I have discussed that O_1 and O_2 in the template cannot both be p-licensed. The constraint is repeated below.

(20) Constraint Between Onsets

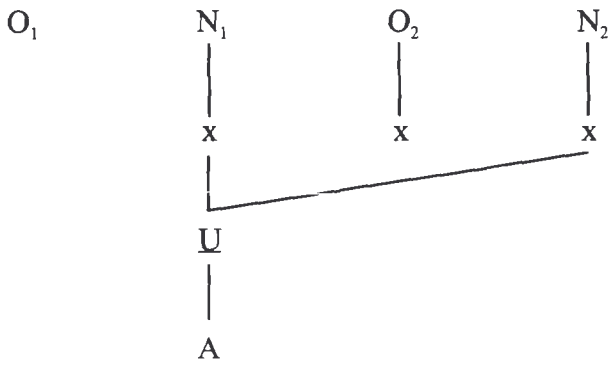
O_1 and O_2 cannot both be p-licensed.

Given the above constraint, structures like (21) will be ruled out in HT, while those in (22) are well-formed.

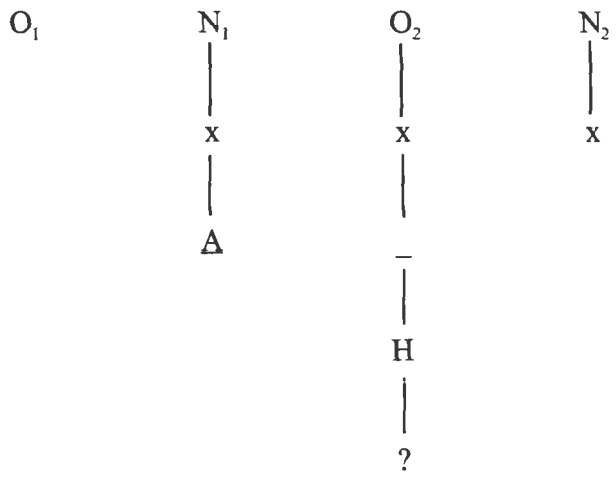
(21) a. *[a:]



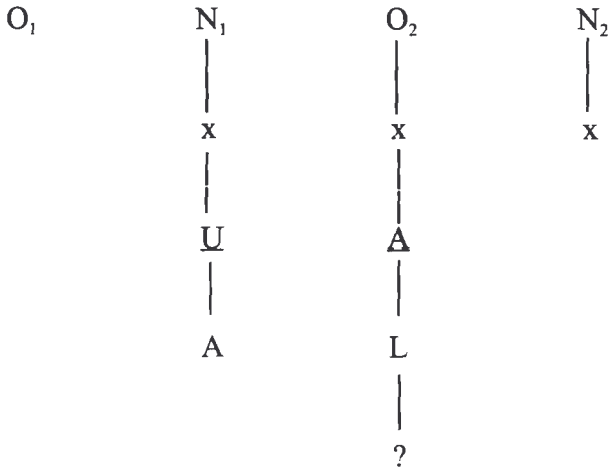
b. *[o:]



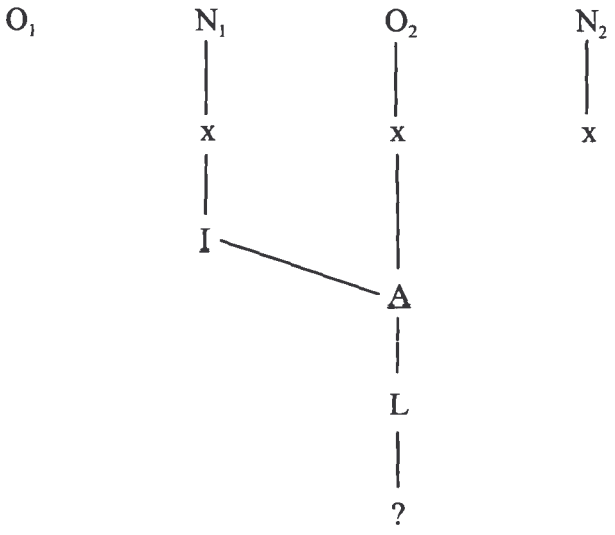
(22) a. [a^{k}] 'chest'



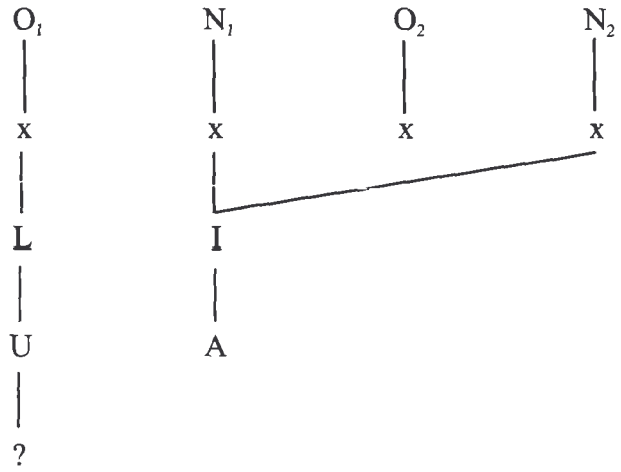
b. [on] 'soft'



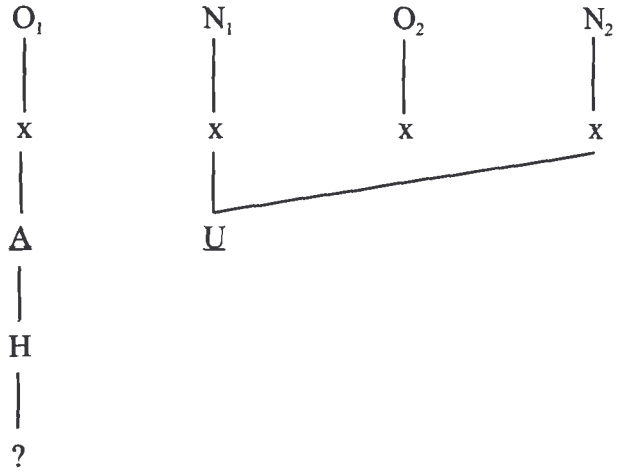
c. [in] 'cigarette'



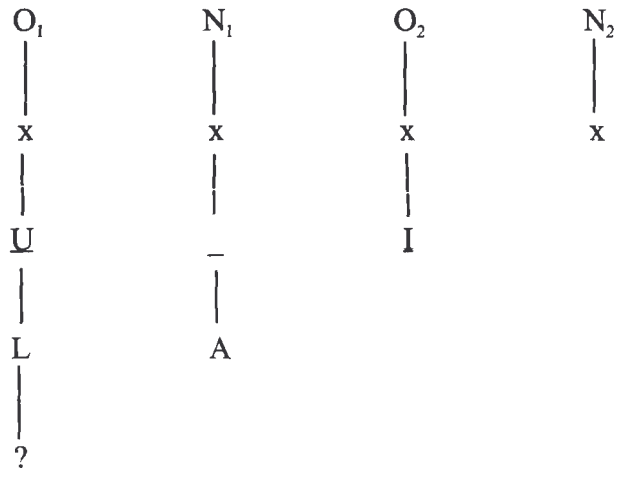
d. [ɓɛ:] 'sheep'



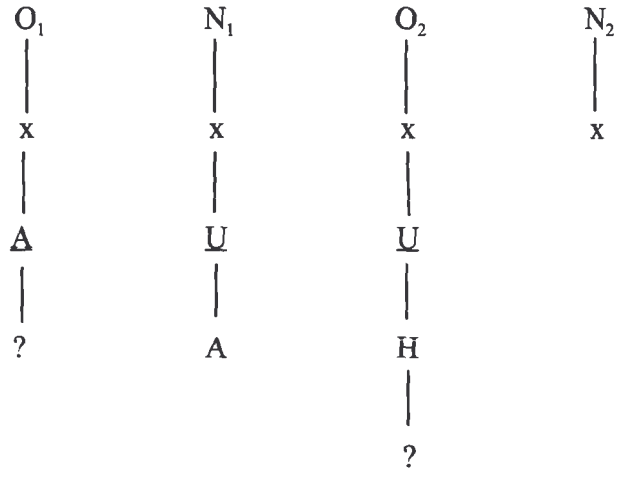
e. [tu:] 'head'



f. [məy] 'to have'



g. [do^{p}] 'to beat'



In (21), structures where the two onsets are both p-licensed are illicit in HT. On the other hand, (22) showed the other possible situations with the two onsets. In (22a) to (22c), the first onset is p-licensed, while O_2 is not. Examples (22d) and

(22e) are just the contrary, namely, the second onset is p-licensed while the first onset is filled. In (22f) and (22g), both of the onsets are filled.

5.4 Summary

This chapter has discussed the interaction between constituents in HT. The discussion of interactions was divided into spreading and constraints. The discovery can be summarized as follows.

1. Interaction between nuclei in HT

   If N_1 is a headed expression, it must spread to N_2 when O_2 is p-licensed.

2. Interaction between O_1 and N_1 in HT

   i. Spreading between O_1 and N_1 in HT

      The element I in N_1 spreads to O_1 when O_2 is filled.

   ii. Constraints between O_1 and N_1 in HT

      (i) The U-Constraint in HT

      If O_1 and N_1 both contain U, then O_1 must be headed.

      (ii) The Empty Nucleus Constraint

      When N_1 is empty, If O_1 contains A, I or U, it must be headed. Otherwise it is headless.

3. Interaction between O_2 and N_1 in HT

i. Spreading between O_2 and N_1 in HT

(i) If U is the head of O_2, it must spread to the head of N_1; If U is in the operator position, the spreading is optional.

(ii) I in O_2 may spread to the head of N_1.

I and U in O_2 may spread to the head of N_1.

ii. Constraints between O_2 and N_1 in HT

When N_1 is an empty nucleus, O_2 cannot contain I or U.

4. Constraint Between Onsets

O_1 and O_2 cannot both be p-licensed.

5. Constraint Between O_2 and N_2

O_2 and N_2 cannot both be filled.

Chapter Six

Tone System[1]

6.0 Introduction

This chapter will deal with the tone system of HT within the framework of the tone theory of GP. Studies on the tone system of some languages have been carried out within the framework GP[2], but the tone system of the Zhuang languages has not been touched upon, not to the mention the specific variety in question. On the other hand, studies on some varieties of the Zhuang languages in the past decades did cover descriptions of tones. However, so far as I know, all of them are descriptive in nature and were carried out on the basis of the "tone values" (see e.g. Wei & Qin 1980, Zhang & Wang 1984, Wang et al. 1979, Li 1986, etc.), which define the tones of all the Chinese languages with five numbers, namely 1 to 5, e.g. [ma] (55) 'Mom' in Mandarin, or [nam] (42) 'water' in Jiangnan To (Qin 1996: 249), etc. The

[1] An earlier draft of this chapter was read at the annual forum of the Linguistic Society of Hong Kong in December 2000. The revised version has been accepted by The First International Conference on Formal Linguistics (China) held in Changsha, China in June 2001.

[2] See e.g. Kaye (2000b), Lin (2001), Chen (2001), etc.

five numbers, 1 to 5, are used to indicate the tone levels in human languages, with 1 being the lowest and 5 the highest tone. This 'tone value' theory has been so widely and uniformly accepted and applied that no criticism or alternative proposals have been made about it for decades. Yet the following will show its inadequacy first and then the rationale of using the tone theory of GP to analyze tone systems.

The fundamental problem with the "tone value" theory is its methodology, namely why are exactly five tone levels defined? Why not four? Or six? Thus, one crucial problem with this framework is that there is neither theoretical nor empirical evidence that the set of numbers reflects human cognitive ability on the acquisition of tones. On the other hand, in a system with five tone levels and tonal patterns consisting of two nuclei, like Mandarin, Cantonese, etc., we will expect 25 possible combinations of tonal patterns. Yet these never occur in any language. In addition, the tone value theory also has trouble expressing tone sandhi with numbers, e.g. why, in Mandarin, a sequence of Tone 3 – Tone 3 yields the sequence Tone 2 – Tone 3. From these three major problems, it is not difficult to see that the so-called "tone values"' are inadequate for analyzing the tone systems of human languages.

Instead of five tone values, there are only two elements for forming tonal expressions in GP, i.e. H(igh) and L(ow), plus an empty expression (_) or the so-called mid-tone or toneless expression. Given the same assumption that each tonal pattern consists of two positions, these three will come up with 9 possibilities of free combination, which is a much more reasonable solution to the tone systems of

human languages. The combinations are as follows.

(1)	(H, H)	(_, H)	(L, H)
	(H, _)	(_, _)	(L, _)
	(H, L)	(_, L)	(L, L)

Of the above nine combinations, the first and the last one—(H, H) and (L, L), are ruled out by one of the universal principles of the tone systems of templatic languages, which is repeated below for the reader's convenience.

(2) Extended OCP: a given tone may appear only once in a tonal pattern. Thus there cannot be two Hs or Ls in a single tonal pattern, but _ _ _ is well formed.

This universal principle reduces the nine combinations in (1) to seven.

(3)	~~(H, H)~~	(_, H)	(L, H)
	(H, _)	(_, _)	(L, _)
	(H, L)	(_, L)	~~(L, L)~~

These are the possible tonal patterns for a phonological system with two nuclei. As mentioned in Chapter Three, GP requires licensing constraints. They regulate the behavior of the three contrasts that are parameterized in specific languages and

provide explanations to why a language can have only some of the tonal patterns.

Let us take Mandarin as an example, the data of which may be more familiar to people. There are four tonal patterns in Mandarin, namely High Level, High Rising, Low Rising and High Falling. Kaye (2000b) proposed the Mandarin Tonal System as follows.

(4) The Mandarin Tonal System

a. Inventory: Complex (H, L or toneless)

b. Spreading: H and L, Obligatory, Left to right

c. Head: Intrinsic (H)

(4a) claims that there are three tonal contrasts on a single nuclear position in Mandarin. (4b) defines the spreading constraint in the tone system of Mandarin. Elements H and L must spread rightward when they are followed by the empty position. (4c) proposes that H is the intrinsic head in Mandarin. From this it follows that every licit tonal pattern in Mandarin must contain the element H, whatever else it may contain. This constraint eliminates all the tonal patterns that do not contain H. The process is shown below.

(5)

~~(H, H)~~	(_, H)	(L, H)
(H, _)	~~(_, _)~~	~~(L, _)~~
(H, L)	~~(_, L)~~	~~(L, L)~~

In (5), the expressions that contain two Hs and two Ls have been eliminated by the universal principle of the Extended OCP. (4c) further rules out the patterns that do not contain H. Thus there are exactly four tonal patterns that satisfy the constraint proposed in (4c). They represent the four tonal patterns in Mandarin.

(6) Mandarin Tonal Patterns

Lexical Forms	**Surface Forms**	**Glossary**
(H, _)	High Level	[ma:] 'mom'
(_, H)	High Rising	[ma:] 'linen'
(L, H)	Low Rising	[ma:] 'horse'
(H, L)	High Falling	[ma:] 'curse'

The above are the four lexical tonal patterns in Mandarin. As (4b) requires that H and L must spread to the empty position on the right, (H, _) becomes (H >> H) after spreading, achieving the tonal pattern of High Level.[3]

Having shown the inadequacy of the traditional "tone value" theory and the rationale of using the tone theory of GP, this chapter will discuss the tone system of HT within the framework of GP. The chapter consists of four sections. The first

[3] L-spreading takes place in the tone sandhi involving "tone 3". As the tone system of Mandarin is not my major concern here, the reader may refer to Kaye (2000b) for details.

section will examine the tone system to which HT belongs—complex or simple. Section 6.2 will discuss the tonal licensing constraints of the language, namely how the tones may combine with each other to derive more tonal patterns, how they spread, their length and the head of tonal patterns. Following these licensing constraints, the lexical tonal patterns, i.e. the specific combinations of the basic tones, of HT will be established. Finally, section 6.3 will summarize the discussion of the chapter.

6.1 Complex Tone Systems

In Chapter One, I noted that GP proposes two main types of tonal inventories: simple and complex. The simple tone system is composed of two levels of tone contrasts, H and toneless. The complex tone systems are broken into two subcategories. One is composed of H, L and toneless. The other is composed of H, L, HL or toneless. Based on the data, HT shows a complex tone system as Mandarin does, i.e. there are three tonal contrasts, H, L and _ on a single nuclear position. Thus there are words like [nap] 'to like', which carries a high tone, [lo:] 'road' of a low tone, and [ða:] 'I', which is toneless. What needs to be mentioned here is that the word for [nap] 'to like' can be both High Level and High Rising, and [lo:] 'road' can be both Low Level and Low Falling. There is no contrast between High Level and High Rising, and between Low Level and Low Falling. This is due to one of the

licensing constraints of the tone system of HT, which will be discussed in detail in the following section, 6.2. There are also words like [ma:] 'horse' and [hay] 'to cultivate', which are High Falling and Low Rising respectively.

There are five contrastive surface tonal patterns in HT. The following minimal pairs of [gaw] illustrate these five tonal patterns.

(7) Tonal Patterns in HT (Provisional Version)

Examples	**Interpretation**	**Surface Forms**	**Provisional Lexical Forms**
[gaw]	'stumble'	Toneless	(_ , _)
[gaw]	'I (fem.)'	High Rising / High Level	(_ , H) / (H, _)
[gaw]	'nine'	Low Level / Low Falling	(_ , L) / (L , _)
[gaw]	'very' / 'old'	Low Rising	(L , _)
[gaw]	'stare'	High Falling	(H , _)

The above list provides examples with the five tonal patterns found in HT. I also presented the provisional lexical expressions for the surface tonal patterns, assuming that, as in Mandarin, H or L spreads to the right empty position. So that (H, _) and (L, _) will become High Level and Low Level respectively. However, it is obvious that there are several problems with the provisional tonal patterns on the assumption of right spreading. First, instead of one exclusive elementary representation, for two of the surface forms, each has two lexical representations.

Secondly, some of the lexical representations are identical. The underlying licensing constraints thus must be discovered so as to define the exact patterns, which is what the next section is going to discuss.

6.2 Licensing Constraints on Tones

Finding out the type of tone system is just the first step in working out the complete tonal system of a language. The second step is to find out the tonal licensing constraints of the language—how the tones may combine with each other to derive tonal patterns, how they spread, how long they can be and the possible head. Bearing this in mind, this section will investigate the rules regulating what a tonal pattern can be—the head of a domain, positions of the components in a domain, if there is spreading, and if so, in which direction. Comparisons will be made between Cantonese/Mandarin and HT.

6.2.1 Tone bearing

In this subsection, the number of nuclei needed for carrying tones will be discussed. Studies on Mandarin have shown that it needs two nuclei in order to have a tonal pattern. Data from HT show the same constraint.

In Mandarin, forms with only one nucleus cannot carry tones, thus particles like [lə], [ma], etc. are always toneless, in contrast to [lə:] 'happy', [ma:] 'curse', etc., which have two nuclei and carry tones. It may then be proposed that, in Mandarin, tones can only appear in a phonological domain and a single nucleus cannot carry tone. In HT, it has the same constraint. Particles like [la], [ni] (as in [gi:n ni] 'here', [va:n ni] 'today' and [nay ni] 'this mroning'), etc. never carry tones. It can thus be concluded that, like in Mandarin, a single nucleus cannot carry tone in HT and tones can only appear in a phonological domain.

6.2.2 Spreading

In the earlier chapters discussing onsets and nuclei, the notion of 'spreading' has been introduced and explained. It refers to the extending of element(s) in one position to a neighboring position. It is a very common phenomenon within and

between onset and nuclear positions, as has been illustrated in previous chapters. It may also take place in the tone system of a language. The following illustrates spreading in the tone system with Cantonese first, then in HT.

6.2.2.1 Spreading in the tone system of Cantonese

Cantonese belongs to the same complex tone system as Mandarin and HT.[4] Thus, three tonal contrasts—H, L and nothing, can be found on a single nucleus. The licensing constraint of the tone system of Cantonese requires that each tonal pattern can have at most one tone. This eliminates the complex expressions in (3) and generates the following tonal patterns in Cantonese (Kaye, 2000b):

(8) Tonal Patterns in Cantonese

a. H _	c. _ L	e. _ _
b. _ H	d. L _	

On the basis of the above tonal patterns, Kaye (2000b) proposed the Cantonese Tonal System shown overleaf.

[4] For more details of the tone system of Cantonese, the reader may refer to Kaye (2000b) and Lin (2001).

(9) The Cantonese Tonal System

a. Inventory: Complex (H, L or toneless)

b. Spreading: H, Optional, Left to right; L, Obligatory unless lexically marked.

c. Head: Positional, Initial

d. LC: A tonal pattern has at most one tone.

(9b) proposes that the spreading direction in the tone system of Cantonese is from left to right. However, H may spread or not spread, while L must spread rightward except for a number of lexically marked forms. H and L in (8a) and (8d) may spread rightward to the empty position and result in a High Level and Low Level respectively. This is exactly what is found in Cantonese. In Cantonese, a form like [maw] 'cat' with the tonal pattern (H, _) may be pronounced either with a level high tone, where the H-spreading takes place, or a High Falling, where spreading does not take place. Both forms are generally acceptable. The structures are shown below.

(10) Spreading: H ⟶ _ ⟹ H H (High Level)

Not spreading: H ⟶̸ _ ⟹ H _ (High Falling)

As for the tonal pattern (8d), there are two possible realizations: Low Rising, if L does not spread rightward, and Low Level if L does spread. Thus there are words like [yaw] 'friend' with Low Rising, where the spreading does not take place, and

[yi:] 'two' with Low Level, where spreading does take place. The structures are as follows.

(11) Spreading: L ⟶ _ ⟹ L L (Low Level)

Not spreading: L ⤫ _ ⟹ L _ (Low Rising)

It is found that the spreading case is much more popular in Cantonese and some non-spreading L _ cases have been eliminated. This indicates that, in Cantonese, L is assumed to spread rightward unless lexically indicated. Having introduced the spreading constraint in Cantonese, now let us turn to the spreading in HT and see if there is any similarity or difference between the two languages.

6.2.2.2 Spreading in the tone system of HT

It has been proposed that, in HT, there are five surface tonal patterns, namely, Toneless, High Level / High Rising, Low Rising, Low Level / Low Falling, and High Falling, the provisional lexical representations being (_ _), (H _) / (_ H), (L _), (L _) / (_ L) and (H_), assuming that there is right spreading of H and L to the empty position as in Mandarin and Cantonese. One obvious problem with these patterns is that there should not be two possible elementary representations for a tonal pattern, even if there is no contrast between patterns like High Level and High Rising. The

assumption here is that there must be one and only one elementary representation for each tonal pattern.

Given the fact that there is no phonological difference between High Level and High Rising and Low Rising and Low Falling. I propose the Spreading Constraint of HT as follows.

(12) Spreading constraint in the tone system of HT

i. H and L may spread leftward to the empty position;

ii. The spreading is optional.

The directionality of spreading goes in the opposite direction of spreading found in Mandarin and Cantonese. In Mandarin and Cantonese, as we have seen, the spreading is from left to right, or Right Spreading. Thus a tonal pattern like (H _) will result in a representation of High Level. In HT, however, Right Spreading is prohibited. So tonal patterns like (H _) and (L _) will not result in High Level and Low Level respectively, as the elements H and L will not spread to the right empty position. Rather they will get High Falling and Low Rising. This rules out (H _) and (L _) as the possible lexical representations for High Level and Low Level. Therefore (_ H) and (_ L) are left as the correct representations for High Level / High Rising and Low Level / Low Falling. When H and L spread to the left, they will result in High Level and Low Level. If they do not spread, they will turn out to be High Rising and Low Falling. The structures and derivations for these tonal

patterns are presented below.

(13) a. H —×→ _ ⟹ H _ (High Falling)

b. L —×→ _ ⟹ L _ (Low Rising)

c. _L
- _ ←— L ⟹ L L (Low Level)
- _ ←×— L ⟹ _L (Low Falling)

d. _H
- _ ←— H ⟹ H H (High Level)
- _ ←×— H ⟹ _H (High Rising)

As (12ii) proposes that the spreading is optional, HT speakers do not distinguish between High Level and High Rising and Low Level and Low Falling. Thus, in (13c) and (13d), (_L) and (_H) each has two possible surface forms. The final version of the tonal patterns in HT is provided overleaf.

(14) Tonal Patterns in HT (Final Version)

Examples	Interpretation	Surface Forms	Lexical Forms
[gaw]	'stumble'	Toneless	(_ , _)
[gaw]	'I (fem.)'	High Rising / High Level	(_ , H)
[gaw]	'nine'	Low Level / Low Falling	(_ , L)
[gaw]	'very' / 'old'	Low Rising	(L , _)
[gaw]	'stare'	High Falling	(H , _)

The spreading constraint in (12) indicates that HT has a different spreading direction from Mandarin and Cantonese. To generalize the language-specific rules regulating the spreading behavior, I propose the following parameter of spreading.

(15) Parameter of the Direction of Spreading in Tone Systems

In a tone system,

i. When the head is intrinsic, the spreading in tonal patterns is directional: [Left to right] / [Right to left];

ii. When the head is positional, the spreading is from the head position.

The underlying assumption of the above constraint is that there can be only one way of spreading in a tone system. In the tone system where the head is intrinsic, like Mandarin, because the head is intrinsic, its position may vary in different tonal

patterns. The spreading must be directional so that it will always take place in the same direction. So in Mandarin, the spreading is always rightward. On the other hand, when the head is positional in a tone system, whether initial or final, the spreading from it will always be of the same direction. From (9), we can see that Cantonese belongs to the second type of spreading. The head in Cantonese tonal system is positional, which is the initial position in a tonal pattern. Thus, the spreading takes place from the left, or the head, position. Having demonstrated the two types of spreading in the tone systems, now let us turn to the next section to see the head of the tonal patterns in HT and decide to which type of spreading HT belongs.

6.2.3 Head of the tonal patterns

From (8) and (14), we can find that Cantonese and HT share the same set of tonal representations. But due to their different spreading directions, they result in different surface patterns. The lexical tonal representations and surface patterns of the two languages are repeated below for the sake of comparison.

(16) Comparison Between the Tonal Patterns in Cantonese and HT

Tonal Patterns	**Surface Forms in Cantonese**	**Surface Forms in HT**
(_ , _)	Toneless	Toneless
(_ , H)	High Rising	High Rising / High Level
(_ , L)	Low Falling	Low Falling / Low Level
(L , _)	Low Level / Low Rising	Low Rising
(H , _)	High Level / High Falling	High Falling

The above table shows how the same set of tonal representations achieve different surface patterns in Cantonese and HT. We have proposed that, in Cantonese tonal system, the spreading always takes place from the left. Thus the initial position in the tonal patterns is proposed to be the head. In the tone system of HT, the spreading always takes place from right to the left. In addition, as no tonal element, either H or L, is required to appear in every tonal pattern, we can conclude that there is no intrinsic head in the HT tonal patterns. I thus propose the following head constraint for the tone system of HT.

(17) Head Constraint in the Tone System of HT

i. The head is positional.

ii. The final position is the head.

From the above constraint, we can conclude that, as in Cantonese, the spreading in the tonal system of HT also belongs to (15ii). It is always from the head. Finally, let us discuss the licensing constraint in the tone system of HT.

6.2.4 Licensing constraint

The tonal patterns in (8) and (16) show a mirror image with the combination of two tones, H and L, and the empty position. It also shows that there can be at most one element in each tonal pattern, thus we can propose that HT shares the same Licensing Constraint of the tone system as Cantonese, which is shown below.

(18) Licensing Constraint in the Tone System of HT

There can be at most one element in each tonal pattern.

This constraint provides an explanation to why tonal patterns with two elements, like (H L) and (L H), which are well-formed in languages like Mandarin, are ruled out in HT.

6.3 Summary

This chapter has been focusing on the tone system of HT, trying to discover the underlying licensing constraints and tonal patterns of HT within the framework of GP. It first discussed problems with the tone value system, a traditional and popular way of doing tone studies, and why GP is more scientific. The discovery in this chapter can be summarized as follows.

1. The HT Tonal System
   - i. Inventory: Complex (H, L or nothing)
   - ii. Spreading: Right to Left; Optional
   - iii. Head: Positional; Final
   - iv. LC: A tonal pattern has at most one tone.

The chapter also made systematic comparisons among Mandarin, Cantonese and HT, which showed that HT shares more similarities with Cantonese. The comparisons also helped to establish the following Parameter of the Direction of Spreading in Tone Systems.

2. Parameter of the Direction of Spreading in Tone Systems

In a tone system,

i. When the head is intrinsic, the spreading in tonal patterns is directional: [Left to right] / [Right to left];

ii. When the head is positional, the spreading is from the head position.

In addition, the chapter discovered that, like in Mandarin, in HT, two nuclei are needed in order to carry tones. Finally, the tonal patterns of HT are concluded as follows.

3. The Tonal Patterns in HT

Tonal Patterns	**Surface Forms**
(_ , _)	Toneless
(_ , H)	High Rising / High Level
(_ , L)	Low Falling / Low Level
(L , _)	Low Rising
(H , _)	High Falling

Conclusion

In this dissertation, I have discussed the phonological system of Hongkui To within the theoretical framework of GP.

First, I proposed the phonological template of HT, which was discovered to be two pairs of onsets and nuclei. However, this is just the basic template of the language, as an augmented template, $O_1N_1O_aN_aO_2N_2$, is allowed to derive from the primary $O_1N_1O_2N_2$. There is also a p-licensing constraint in this template, which requires that N_2 has no lexical material.

Through working on the nuclear system of HT, it was found that there are seven lexical nuclei in this language. The licensing constraints, based on which the nuclear system is established, are that I and U must always be head and they cannot combine. Furthermore, it was proved again that A cannot to be a licensor in the nuclear system, a constraint which has been found in many languages. There are also constraints on spreading, which requires that headless expressions may not spread. It is also found that, in HT, I must spread. If it cannot spread to the right, it spreads to the left.

In considering the onset system and phonological representations of onsets, I proposed the U-Constraint and the Empty Nucleus Constraint in HT, based on which the phonological representations of onsets are suggested.

I have also presented a specific chapter for the discussion of inter-constituent interaction. So far, there have been found two main types of interactions between constituents—interaction between nuclei and interaction between onset and nucleus. Spreading was the interaction discovered between the two nuclei. Interaction between onset and nucleus, on the other hand, is far more complex than that between nuclei, as there are two onsets involved. The constraints that O_1-N_1 interaction and O_2-N_1 interaction share are leftward spreading. The specific constraints in interactions were also proposed.

I also discussed the tone system of HT, as it is a "tone language". The language was discovered to have a complex tone system, viz. there are three tonal elements—H, L and "_", at the lexical level. To derive all and only the five tonal patterns of HT from these three elements, I presented the underlying constraints. In way of spreading, it shares the same direction as in the onset system, i.e. from right to left. It was also discovered that in each tonal pattern, there can be at most one tonal element. The head of the tonal pattern was found to be positional, which is the final position.

Although this dissertation is the first phonological study on HT, it shows that GP is the appropriate theory that can provide adequate explanations for various aspects of phonology of this language and related phonological phenomena and behavior. However, as we know, science must develop and perfect itself. So must GP. I am fully aware that this dissertation is just a beginning to our understanding of HT phonology and Zhuang phonology, to a larger extent. I hope that it has its own contribution and position in the development of this field.

Bibliography

Backley, P. & T. Takahashi (1998). Element activation. In Cyran, E. (eds.) *PASE Studies and Monographs (Volume 4): Structure and Interpretation—Studies in Phonology*. Wydawnictwo Lublin, Poland.

Charette, M. (1988). *Some Constraints on Governing Relations in Phonology*. Ph.D. dissertation. Mcgill University.

Charette, M. (1990). License to govern. *Phonology* **7.2**, 233-254.

Charette, M. (1991). *Conditions on Phonological Government*. Cambridge University Press, Cambridge.

Charette, M. & Göksel, A. (1994). Switching and vowel harmony. *SOAS Working Papers in Linguistics and Phonetics* **4**. 71-78

Charette, M. & Göksel, A. (1996). Licensing constraints and vowel harmony in Turkic languages, Manuscript, SOAS, University of London.

Charette, M. & Göksel, A. (1998). Licensing constraints and vowel harmony in Turkic languages. In Eugeniusz Cyran (ed.), *Structure and Interpretation: Studies in Phonology*, 65-88. Lublin: Wydawinctwo.

Chen, J. [陈洁] (2001). MA thesis, Guangdong University of Foreign Studies

Cheng, C-C [郑锦全] (1973). *A synchronic phonology of Mandarin Chinese*. The Hague, Mouton.

Chiu, A. D-L (1992). *Syllable structure of Taiwanese in the Framework of Government Phonology*, MA thesis, SOAS, University of London.

Cobb, M. (1995). Licensing constraints and vowel harmony in Uyghur. *SOAS Working Papers in Linguistics and Phonetics* **3**. 40-64.

Duanmu, S. [端木三] (1990). *A formal study of syllable, tone, stress and domain in Chinese languages*, PhD dissertation, Massachusetts Institute of Technology.

Goh, Y-S [吴英成] (1996a). *The Segmental Phonology of Beijing Mandarin*. PhD dissertation, SOAS, University of London.

Goh, Y-S [吴英成] (1996b). *北京方言音韵成素研究* (*Beijing Fangyan Yinyun Chengsu Yanjiu*) [The Segmental Phonology of Beijing Mandarin]. Taipei: WenHe Publication Ltd.[文鹤出版有限公司].

Goldsmith, J. A. (1990). Augosegmental and metrical phonology. Oxford: Blackwell.

Guangxiqu Yuwei Yanjiushi [广西区语委研究室 编] (1994). *壮语方言土语音系（Zhuangyu Fangyan Tuyu Yinxi）*[The Phonology of the varieties of To of Zhuang Language] (in Chinese), 广西民族出版社出版发行。

Harris, J. (1990). Segmental complexity and phonological government. *Phonology*, **7.1**, 255-300.

Harris, J (1992). Licensing inheritance. *UCL Working Papers in Linguistics*, **4**, 359-406.

Harris, J. (1994). *English Sound Structure*, Oxford: Blackwell Publishers.

Harris, J. (1997). Licensing inheritance: an integrated theory of neutralization. *Phonology*, **14**, 315-370.

Harris, J. & J. Kaye (1990). A tale of two cities: London glottaling and New York City tapping. *The Linguistic Review* **7-3**. 251-274.

Kaye, J. (1990). Coda Licensing. *Phonology* **7.2**, 301-330.

Kaye, J. (1995). Derivations and interfaces. In Durance, J. & F. Karamba (eds.) *Frontiers of Phonology*. Longman Group UK Limited.

Kaye, J. (1999a). Biological foundation of linguistics. Inaugural speech of special professor. Guangdong University of Foreign Studies.

Kaye, J. (1999b). A short theory about tones. Paper presented at '99 LSHK Annual Research Forum, Hong Kong.

Kaye, J. (2000a). Working with licensing constraints. Manuscript at Guangdong University of Foreign Studies.

Kaye, J. (2000b). A short theory about tones. Manuscript at Guangdong University of Foreign Studies.

Kaye, J., J. Lowenstamm, & J-J Vergnuad, J. (1985). The internal structure of phonological elements: A theory of charm and government. *Phonology Yearbook*, **2**, 305-328.

Kaye, J., J. Lowenstamm, & J-J Vergnaud (1990). Constituent structure and government in phonology. *Phonology* **7.2**, 193-231.

Kim, S-J (1996). *The Representations of Korean Phonological Expressions and Their Consequences*. PhD dissertation, SOAS, University of London.

Leben, W. (1973). *Suprasegmental Phonology*, Ph.D. dissertation, MIT, Cambridge.

Li, F-K. [李方桂] (1954). Consonant Clusters in Tai. *Language* **30**, 368-379

Li, F-K. [李方桂] (1984). 汉语和台语 (Hanyu He Taiyu) [Han languages and Tai languages] (in Chinese). *民族语文研究情报资料集* （*Minzu Yuwen Yanjiu Qingbao Ziliaoji*）[Collection of information on the studies of minority languages], 中国社会科学院民族研究所语言室编。

Li, F-K. [李方桂] (1986). 原始台语的声调系统 (Yuanshi Taiyu De Shengdiao Xitong) [Ancient Tai's tone system] (in Chinese). *民族语文研究情报资料集*（*Minzu Yuwen Yanjiu Qingbao Ziliaoji*）[Collection of information on the studies of minority languages], 中国社会科学院民族研究所语言室编。

Li, F-K. [李方桂] (1987). 台语的齿音复辅音 (Taiyu De Chiyin Fufuyin) [The dental consonant clutsters of Tai languages] (in Chinese). *民族语文研究情报资料集* （*Minzu Yuwen Yanjiu Qingbao Ziliaoji*）[Collection of information on the studies of minority languages], 中国社会科学院民族研究所语言室编。

Liang, M. [梁敏] (1982). 壮语形容词、名词、动词后附音节的研究 [Studies on the attached syllables of adjectives, nouns and verbs in Zhuang] (in Chinese). *民族语文研究文集* （*Minzu Yuwen Yanjiu Wenji*） [Collection of studies on minority languages], 青海民族出版社。

Liang, M. & Zhang, J-R [梁敏，张均如] (1996). *侗台语族概论* (*Tongtai Yuzu Gailun*) [An outline of Tai linguistic family] (in Chinese), 中国社会科学出版社。

Lin, Q-M [林秋茗] (2001). MA thesis, Guangdong University of Foreign Studies.

Lin, Y-H [林燕慧] (1989). *Autosegmental treatment of segmental processes in Chinese Phonology*. PhD dissertation, University of Texas at Austin.

Pan, N. [潘宁] (2000). MA thesis, Guangdong University of Foreign Studies.

Qin, G-S. [覃国生] (1996). *壮语方言概论* (*Zhuangyu Fangyan Gailun*) [An outline of Zhuang dialects] (in Chinese) (in Chinese), 广西民族出版社。

Wang, J. [王均], Liang, M. [梁敏], Wei, Q-W. [韦庆稳], Qin G-S. [覃国生] (1979). *壮语及壮汉人民怎样互学语言* （*Zhuangyu Ji Zhuanghan Renmin Zenyang Huxue Yuyan*）[Zhuang languages and how Zhuang and Han people learn the other language] (in Chinese), 民族出版社。

Wei, Q-W. [韦庆稳], & Qin, G-S. [覃国生] (1980). *壮语简志* （*Zhuangyu Jianzhi*）[Sketch of Zhuang languages] (in Chinese), 民族出版社。

Xu, Z. [许卓] (2000). *Phonological template of Hongkui To*, PhD proposal defense, Guangdong University of Foreign Studies.

Yin, Y-M [殷允美] (1989). Phonological aspects of word formation in Mandarin Chinese, PhD dissertation, University of Texas at Austin.

Yoshida, S. (1990). *Some Aspects of Governing Relations in Japanese Phonology*, Ph.D. dissertation, SOAS, University of London.

Zhang, Y-S., & Wang, W. (1984). 壮语和汉语关系初探 (Zhuangyu He Hanyu Guanxi Chutan) [A primary study on the relationship between Zhuang and Han] (in Chinese) in *民族语文论丛*(*Minzu Yuwen Luncong*) [Collection of research on minority languages], 中央民族学院，少数民族语言研究所编。

Zhongguo Shehui Kexueyuan Minzu Yanjiusuo & Guojia Minzu Shiwu Weiyuanhui Wenhua Xuanchuansi [中国社会科学院民族研究所，国家民族事务委员会文化宣传司 主编] (1994). *中国少数民族语言使用情况* (*Zhonguo Shaoshu Minzu Yuyan Shiyong Qingkuang*) [The situation of the usage of Chinese minority languages] (in Chinese), 中国藏学出版社。