VIETNAMESE TONE: TONE IS NOT PITCH

by

Hoa T. Pham

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Linguistics
University of Toronto

© Copyright by Hoa Pham 2001
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
Vietnamese Tone: Tone is not Pitch

Doctor of Philosophy, 2001
Hoa Pham
Graduate Department of Linguistics
University of Toronto

ABSTRACT

This thesis is an in-depth study of the phonology and phonetics of tones in Northern Vietnamese dialects. It examines tone within the framework of markedness theory, the phonetic realizations of tones, the relationship between the phonetics and phonology of tones, and the implications of the proposed analysis for other dialects of Vietnamese.

First, I propose a phonological model in which tones are structurally organized. In this model, tones are grouped together as constituents according to their markedness relationship. Evidence for the structural organization of tones comes from neutralization in dialects and reduplication, and also from tonal frequency in the language, poetry, borrowings and clitics. This evidence points to tones having two major constituents, contour and register, each of which can pattern independently.

Second, I show that the proposed features are grounded phonetically. I argue that instead of pitch height being contrastive as is generally assumed, it is phonation types of creakiness and breathiness which are distinctive as the register feature in North Vietnamese, and the differences in pitch heights are predictable. In the Vietnamese literature, it is claimed that there is a mismatch between the phonetics
and phonology of two tones, boi and nga in reduplication. Flip-flop rules were suggested to account for the mismatch. I show that using phonation types as distinctive feature not only can explain various tonal patterns in the language, it also makes this mismatch an illusory one, which, in turns, makes ad hoc flip-flop rules in previous models unnecessary.

In addition, I argue that the domain of Vietnamese tone is not the syllable or mora, but the rhyme, and that the crucial portion for distinguishing the tones is the middle point in boi and nga, and the endpoint in other tones. Finally, I show that the model correctly predicts certain types of neutralization in other Vietnamese dialects, in language acquisition, and in other tone languages.
ACKNOWLEDGMENTS

I was very lucky to work under the supervision of Keren Rice. I appreciate her devotion of time, energy and thoughtful insights to my work. What made me deeply grateful to her was that among my early incoherent ideas and broken data, she spotted a good seed, picked it out and gave it a great care until it blossomed. Nor could the thesis have been written in time without her incredibly fast feedback.

I am also grateful to all other members of my thesis committee: Hank Rogers for his great support, both academic and technical; Ron Smyth for his help with the pilot study and useful comments later, and his arrangement for a lovely French restaurant lunch after the defense; Peter Avery for helping me from the very early stage of the thesis; Parth Bhatt for spending time to discuss my experiment with special enthusiasm. Peter and Ron also did a wonderful editing job. I am also grateful to my external examiner, Moira Yip, for her fruitful comments. I owe Bill Idsardi for his valuable suggestions, especially with the experiment. I would also like to thank all consultants who were very generous in giving me their time.

My sincere thanks also go to Diane Massam for being always so comfortable to talk to, and for giving me tips for the defense; to Jack Chambers for the firm encouragement I could always rely on.

The financial support for my research came from various sources: Ontario Graduate Scholarship, SSHRC Grant (752-2000-1577), School of Graduate Travel Grant, David Chu Travel Fellowship, General Research Grant from the linguistics department, SSHRC Grant to Professors Rice and Dresher for the Phonology Project, and the University of Toronto Foundation Graduate Award.

iv
I would not have survived the program without my friends. I cannot name them all, but am grateful for all the support I have received: an encouraging smile in the computer room, a dance in a social club, a pointer toward a possibly related article, a comforting shoulder, and a take-a-break-trip when desperately needed.

I thank my family -big and small- for their love. Finally, my deepest gratitude to my late parents for giving me life and nourishing my spirit.
# TABLE OF CONTENTS

| ABSTRACT | ii |
| ACKNOWLEDGMENTS | iv |

## Chapter 1  INTRODUCTION

1.0. Introduction  
1.1. Questions addressed  
1.2. Language background  
  1.2.1. Language classification  
  1.2.2. Dialectal classification  
  1.2.3. Syllable pattern  
1.3. Theoretical assumptions  
  1.3.1. Feature geometry  
  1.3.2. Markedness and complexities  
1.4. Outline of the thesis  

## Chapter 2  PHONOLOGY OF TONES: MARKEDNESS AND COMPLEXITY

2.1. Background: Tone distribution and phonetic descriptions  
2.2. Proposal: a structural representation  
  2.2.1. Tonal features: contour and register  
  2.2.2. Representation of VN tones  
  2.2.3. Summary  
2.3. Evidence for constituents and markedness relations  
  2.3.1. Reduplication and tone harmony: evidence for the independence of register and contour  
  2.3.2. Reduplication: evidence for the markedness of ngang and huyen and for register being independent of contour

vi
| 2.3.3. Frequency of occurrence | 38 |
| 2.3.4. Poetry | 39 |
| 2.3.5. Neutralization in dialects | 42 |
| 2.3.6. Clitics | 47 |
| 2.3.7. Borrowings | 50 |
| 2.3.8. Summary | 53 |
| 2.4. Traditional classification and representation of tones: the mismatch problem | 54 |
| 2.4.1. Representation of Vietnamese tones in the traditional literature | 54 |
| 2.4.2. Mismatch between the phonetics and phonology in reduplication | 62 |
| 2.5. Summary | 67 |

**Chapter 3**  
AN ACOUSTIC STUDY OF THE TONES OF THE HANOI DIALECT

| 3.1. Literature: acoustic studies | 69 |
| 3.2. Experimental design | 79 |
| 3.2.1. Subjects and data | 79 |
| 3.2.2. Procedures and methodology | 80 |
| 3.3. Results | 82 |
| 3.3.1. Descriptions of the eight tones | 82 |
| 3.3.1.1. Fundamental frequency | 82 |
| 3.3.1.2. Remarks on tonal shape | 93 |
| 3.3.2. Spectrographic analysis: waveform and spectrogram | 101 |
| 3.3.2.1. Background in voice qualities (modal, creaky, breathy) | 101 |
| 3.3.2.2. Waveforms and spectrograms | 104 |
| 3.3.2.3. Discussion | 114 |
| 3.3.2.4. Summary | 116 |
| 3.3.3. Other features: height, length, and linear portion | 117 |
| 3.3.3.1. There is no difference in tonal height | 117 |
| 3.3.3.2. Variations in tonal length | 121 |
| 3.3.3.3. Tone linear portion | 129 |
| 3.4. Summary | 136 |
5.1.3. Summary

5.2. The mapping between the phonological representation and the phonetic realization

5.2.1. Phonation types: motivation for a register feature

5.2.2. Mapping between the phonetics and phonology of tone

5.2.3. Neutralization in hoi

5.2.3.1. Neutralization of Contour in hoi

5.2.3.2. Neutralization of Register in hoi in Southern dialects

5.3. Summary

Chapter 6  DOMAIN OF TONE

6.1. Syllable structure in Vietnamese

6.1.1. Tone is independent from the segment

6.1.2. Independence of initials: evidence for the rhyme

6.1.3. Feature sharing: nucleus vs rhyme

6.2. Domain of tone: syllable, rhyme, nucleus, mora

6.2.1. The syllable is not a domain: Initials are outside of the domain

6.2.2. The mora is not the domain of tone

6.2.3. The nucleus is not the domain of tone

6.2.3.1. Tone is distributed over the whole rhyme

6.2.3.2. Length in Vietnamese rhymes

6.3. Conclusion

Chapter 7  SUMMARY AND IMPLICATIONS

7.1. Summary of major findings

7.2. Predictions and implications

7.2.1. Other dialects of Vietnamese
7.2.2. Language acquisition 287
7.2.3. Neutralization in other languages 288
7.2.4. Tonal distribution in other languages 390
7.3. Questions addressed for future research 293
  7.3.1. A mismatch in Chaozhou tones 293
  7.3.2. What are underlying features in all dialects of Vietnamese?
      Implications for perception and typology 295
7.4. Conclusion 301

REFERENCES 302
APPENDICES
Appendix 1 - Map 308
Appendix 2 - Initials 309
Appendix 3 - Subjects' information 310
Appendix 4 - Sentence list 311
Appendix 5 - Pitch of others 318
Appendix 6 - Spectrograms of others 326
Chapter 1

INTRODUCTION

1.0. Introduction

This thesis investigates Vietnamese tones, their markedness relationships, the organization of tones by features, and the nature of these features.

In the literature on tone languages, there is a very common view that assumes that pitch height is the primary, and sometimes only, phonetic correlate of tone. This view creates a problem in Vietnamese, namely there is a mismatch between the phonetic realization of tone and the phonological patterning of tone in Vietnamese reduplication. My thesis questions the assumption that pitch height is the primary phonetic correlate of tone. In my thesis, the term 'tone' is used to refer to a phonological category, something contrastive that is not, in fact, pitch height. The term 'pitch height' or 'tonal height' refers to the fundamental frequency (F0), one of the acoustic properties employed in the recognition of tone. In particular, I argue that, in Vietnamese, pitch height is not a basic tonal feature; therefore, pitch register is not an adequate phonological feature to represent Vietnamese tones. Rather the laryngeal features of creakiness and breathiness are primary in signaling tone. Pitch height is derived from these features and from features describing tonal shape.

The complexities of Vietnamese tones offer fertile ground to study the nature of tone. Vietnamese has several properties that make a study of its tones especially fruitful.
First, in Vietnamese, tones are lexical, and every syllable must bear a tone. Second, tone is not affected by neighbouring tones in any environment - there is no tone sandhi. Finally, the tonal inventory of Vietnamese is rich, with either six or eight contrastive tones depending on one’s analysis.

1.1. Questions addressed

In this thesis I address the following questions concerning tonal features and the structure of tones. What are the features of tones in Vietnamese? In a language with a rich tonal inventory such as Vietnamese, are tones organized in some sort of structure? If they are, what is the internal structure of tonal features? What are the phonetic features of tones? How close are phonological features and phonetic features in Vietnamese tones?

Consider first both the phonological and the phonetic properties of tones. In the Vietnamese linguistic literature the question of why a tone patterns in a particular way is seldom posed. For instance, in one dialect tone A neutralizes to tone B: why not B to A, or A to C? Tonal features are not clearly defined based on phonological patterning; rather, tonal features are determined by their phonetic properties. More importantly, when tonal features are discussed, the relationship between those features has never been addressed. Rather, tones have been represented as a bundle of unorganized features (Cu et al. 1977, Thompson 1965, Doan 1977, Vo 1997, Alves 1997 and others) such as [high], [tense], [flat], [glottalic].

Next consider the markedness of tonal features. While the markedness of segmental features is often discussed in the phonological literature (e.g., Paradis and
Prunet 1989, Avery and Rice 1989, Rice 1992, Archangeli 1988), there is little discussion of the markedness of tones in the literature in large tone systems. The question of how to define the complexity of tone is almost neglected.

Third, consider the phonetics of tones. In the Vietnamese literature, the assumption is made that the primary phonetic correlate of tone is pitch height. However, tones in Vietnamese also have laryngeal realizations of breathiness and creakiness. A detailed study of the phonetic constants associated with tones is not available.

This study has two major goals. First, I propose a hierarchical structural representation of Vietnamese tones that captures their markedness relations. The model accounts for the patterning of tone in reduplication and for several types of tonal neutralization found in the language. Second, by examining the acoustics of Vietnamese tones, I provide phonetic evidence that the phonation properties of creakiness and breathiness are important phonological features of Vietnamese tones. I also demonstrate how tonal features are realized phonetically. This is the first time in the Vietnamese literature that both the phonetics and the phonology of tones have been unified in one model.

1.2. Language background

In this section I give some background on the Vietnamese language and its syllable types, and I introduce the classification of dialects with respect to tones.
1.2.1. Language classification

Vietnamese is a Viet-Muong language in the Mon-Khmer group, and belongs to the Austroasiatic family. Viet-Muong is one of nine subgroups in the Mon-Khmer group. Among the approximately 80 million speakers of Austroasiatic languages, 65 million people spoke Vietnamese in 1988 (Tran 1999:131). It is now spoken by about 78 million people in Vietnam (Dodd and Lewis 1998), and by about two million speakers living all over the world, with half of this number in North America.

1.2.2. Dialect classification

The dialects of Vietnamese can be classified into several different major groups depending on the aspect of the grammar under consideration. For instance, based on final consonants, Vietnamese can be divided into two major groups, North Vietnamese and South Vietnamese (Gordina and Bystrov 1970 cited from Hoang Thi Chau 1989, Pham 1998). With respect to tones, Vu 1982, Hoang Cao Cuong 1989 and Hoang Thi Chau 1989 agree that there are three major dialect groups: North Vietnamese, Central Vietnamese and South Vietnamese (see the map in Appendix 1). Northern dialects have the largest tonal inventory and are the major focus of my study. Hanoi, one of Northern dialects, is the standard dialect.

1.2.3. Syllable pattern

The overall syllable pattern in all Vietnamese dialects is given in (1).

(1) \((C(w))V(G, N \text{ or } T)\)
I focus on the final segments; therefore, the initial consonant is simplified as C. The initial consonant inventory of Northern dialects is given in Appendix 2. C stands for any consonant elements. V can be a short or long vowel or a diphthong with an on-glide. G represents an off-glide, N stands for a nasal consonant, and T represents a voiceless stop. Each syllable must bear a tone. The domain of tone is discussed in Chapter Six. Besides tone, only V is obligatory in the syllable.

While any consonant can occur in initial position, only a limited number of consonants can occur in final position. The manner of final consonants is important to this study, and thus the full inventory of finals given in (2).

(2) Possible syllable-final segments (phonetic)

<table>
<thead>
<tr>
<th></th>
<th>labial</th>
<th>alveolar</th>
<th>palatal</th>
<th>labio-velar</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obstruents</td>
<td>p</td>
<td>t</td>
<td>c</td>
<td>kp</td>
<td>k</td>
</tr>
<tr>
<td>Nasals</td>
<td>m</td>
<td>n</td>
<td>ĵ</td>
<td>ɲm</td>
<td>ɲ</td>
</tr>
<tr>
<td>Glides</td>
<td>w</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The inventory in (2) consists of nasals, obstruents and two glides. All final obstruents are voiceless stops. Stops are unaspirated or inaudibly released. Palatal, labio-velar and plain velar are allophones, determined by the quality of the preceding vowel.¹

¹ Pham (1998) argues that there is a placeless consonant in final position in all Vietnamese dialects. Place of
1.3. Theoretical assumptions

1.3.1. Feature geometry

Following Clements 1985 and McCarthy 1988, among others, I assume that features are organized in a hierarchical configuration. The hierarchical structure allows a set of features to undergo certain phonological processes without affecting features of another group. Elements that pattern together as a natural class are dominated by the same organizing node. The working model is illustrated in (3). R is a root node, X and Y are organizing nodes, z and p are content nodes.

(3)  \[
\begin{array}{c}
R \\
\downarrow \\
X \quad Y \\
| \quad | \\
z \quad p
\end{array}
\]

The sisterhood relationship between the two organizing nodes X and Y in (3) allows a process to affect X without affecting Y.

1.3.2. Markedness and complexity

I assume first that markedness can be expressed by the way in which the relative complexity of the system is encoded (Archangeli 1988). Following the work of Steriade 1987, Mester and Itô 1989 and Avery and Rice 1989, I assume the theory of Contrastive Specification. In the particular version of this theory that I adopt (after Avery and Rice articulation, however, is not a concern in this paper. Manner of final consonants is discussed in Chapter Four.
all features are monovalent. A feature must be present underlyingly if it is contrastive. Non-contrastive or unmarked features, which play no role in the phonology of a language, are absent from underlying representations. Moreover, while the unmarked features are generally absent underlyingly, organizing nodes are, I assume following Rice 1993, present in underlying representation if they define a property inherent in the segment.

A feature must be present underlyingly if it produces contrasts in the system. Markedness can be overridden by contrast within an inventory. In particular, I follow Rice & Avery 1993 in assuming that structure is elaborated under pressure from the phonology to contrast two sounds. As contrasts are introduced, more complex structures are added to the existing less complex ones already in an inventory. The more marked a segment, the more structure it has (Avery & Rice 1989, Rice & Avery 1991, Kaye, Lowenstamm & Vergnaud 1985, among others). The complexity of structural representations, therefore, increases as the number of contrasts in the inventory increases. Note that complexity has a relative meaning. A complex representation is one that is relatively more complex than another representation.

Following Drescher & van der Hulst 1998 I adopt the assumption that there are two types of complexity: local and non-local. A node C has local complexity if it branches while others do not, as in (4a), or if C has an immediate dependent when others do not, as in (4b).
(4) Local complexity

\[
\begin{array}{ccc}
\text{a. complex} & \text{simple} & \text{b. complex} \\
C & C & C \\
\downarrow & | & | \\
D & E & D & D
\end{array}
\]

The second type is non-local complexity, which concerns the internal structure of the daughters of a node. For example, in (5), node C has non-local complexity if C dominates D and D dominates E and F, as in (5a). It is simple if C dominates D and D dominates E but D does not branch, as in (5b).

(5) Non-local complexity

\[
\begin{array}{ccc}
\text{a. complex} & \text{simple} & \text{b. simple} \\
C & C & C \\
| & | & | \\
D & D & | \\
\downarrow & | & | \\
E & F & E
\end{array}
\]

Within non-local complexity, if we assume that a structure with more specified features in a hierarchical system, either vertically or horizontally, should count as more complex than a structure with fewer specified features, then (6a) is vertically more complex than (6b). The only difference between the two is that (6a) has an extra dependent feature G under F.
(6) Vertical complexity

a. complex
   \[ \begin{array}{c}
   \text{C} \\
   \downarrow \\
   \text{D} \\
   \downarrow \\
   \text{F} \\
   \downarrow \\
   \text{G}
   \end{array} \]

b. simple
   \[ \begin{array}{c}
   \text{C} \\
   \downarrow \\
   \text{D} \\
   \downarrow \\
   \text{F} \\
   \downarrow \\
   \text{G}
   \end{array} \]

In (7), (7a) is horizontally more complex than (7b). The only difference between the two is that in (7a), D and E both have dependents while in (7b) only D has an independent.

(7) Horizontal complexity

a. complex
   \[ \begin{array}{c}
   \text{C} \\
   \downarrow \\
   \text{D} \\
   \downarrow \\
   \text{F} \\
   \downarrow \\
   \text{G}
   \end{array} \]

b. simple
   \[ \begin{array}{c}
   \text{C} \\
   \downarrow \\
   \text{D} \\
   \downarrow \\
   \text{F}
   \end{array} \]

I also assume that for an underspecified feature, default rules can act as a phonetic implementation component (Rice and Avery 1989). A delinking rule serves to delink the content nodes in a neutralization environment. In Vietnamese, delinking,
either of the laryngeal or contour features of tone, occurs to yield unmarked features or structures. For instance, a more complex tone, thus more marked, is reduced to the less complex, or less marked, one.

1.4. Outline of the thesis

The remainder of the thesis is organized as follows.

Chapter Two presents the proposal for the structure of Vietnamese tones based on markedness, and phonological evidence for the structure. Markedness of tones is based on their complexity, which is evaluated according to tonal patterning. Evidence for markedness of tones is drawn mainly from tonal neutralization. This chapter analyzes the six tones that occur in the same environment, namely sonorant-final syllables. This chapter also introduces the classification of tones in the traditional literature and the problem it creates.

Chapter Three deals with the phonetics of tones. I review the acoustic experiments in the literature before presenting an experimental study of tones that I conducted with speakers of the Northern dialect. In particular, I look at various aspects of tones: fundamental frequency, shape, length, linear portion, and phonation types such as creakiness and breathiness. One of the most important findings is the instability of F0. The F0 of pitch varies from speaker to speaker, from token to token, and from form to form. However, phonation types are the most stable feature across speakers and tokens. Another finding is that the endpoint of a tone is the most crucial portion of a tone for differentiating that tone from others, except in curved tones where the middle part is crucial.
Chapter Four discusses the size of the tonal inventory, either six tones or eight tones. I present evidence to support the eight-tone hypothesis. I also present arguments about the representation and markedness of the two tones that occur in obstruent-final syllables.

Chapter Five presents the implications of the phonetic results for the phonology of tones. I argue that phonation type is the primary phonetic correlate of tones in Vietnamese. The pitch height is not constant and is, I argue, predictable on the basis of phonation type. By using the phonation features of creakiness and breathiness as laryngeal features, I argue that the mismatch between the phonetics and phonology of tone in reduplication that is claimed in the literature, is an illusory one. I also discuss how the phonetics of tones is realized and how the phonology of tones maps to the phonetics.

Chapter Six discusses the domain of tones. I argue that in Vietnamese, the domain of tone is not the syllable, nucleus or mora. It is smaller than a syllable and larger than a nucleus; it is the rhyme.

Chapter Seven summarizes the major findings of the thesis. I introduce some similar problems in other tonal languages, and discuss implications of the proposed model for other dialects and languages. I also address some questions about perception and typology raised by my thesis for further research.
Chapter 2

PHONOLOGY OF TONES: COMPLEXITY AND MARKEDNESS

In this chapter I examine the phonology of Vietnamese tones, focusing on their structural representations and their markedness with respect to one another. Evidence for the markedness of different tones is drawn mainly from tonal neutralization in various phonological domains.

Section 2.1 provides some background on the distribution of tones and their phonetic descriptions. Section 2.2 proposes a hierarchical structure for tones, followed by evidence for the structure in Section 2.3. Section 2.4 reviews tonal representations in the traditional literature and identifies some of the problems with the traditional accounts.

2.1. Background: Tone distribution and phonetic descriptions

In this section I provide some background on the distribution of tones in Vietnamese and give a brief phonetic description of tones in order to give the reader necessary information about two important properties of tones, pitch and shape. The phonetics of tone is studied in greater detail in Chapters Three and Five.
2.1.1. Tone distribution

Vietnamese has eight tones (see Chapter Four) ngang, huyen, sac1, nang1, hoi, nga, sac2 and nang2. The distribution of these tones is shown in (8). In open or sonorant-final syllables, ngang, huyen, sac1, nang1, hoi and nga occur. In stop-closed syllables, only sac2 and nang2 occur. I call the tones known as sac and nang in non-stop final syllables sac1 and nang1. In stop-closed syllables, I call them sac2 and nang2. See Chapter Three for phonetic details about these tones, and Chapter Four for their phonological status.

(8) Tone distribution

<table>
<thead>
<tr>
<th>Syllable types</th>
<th>e.g.</th>
<th>ngang</th>
<th>huyen</th>
<th>sac1</th>
<th>nang1</th>
<th>hoi</th>
<th>nga</th>
<th>sac2</th>
<th>nang2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV</td>
<td>ba</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVG</td>
<td>ba:w</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVN</td>
<td>l'm</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVT</td>
<td>ba:t</td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.1.2. Some background on the phonetic properties of tones

A set of minimal pairs based on tones is given in (9). The Vietnamese name of each tone is next to each item. The Vietnamese orthography is in the third column. Note that in the orthography, only six diacritics are given. Sac1 and sac2 share the same diacritic. Similarly, nang1 and nang2 have the same diacritic.

---

2 I will use the Vietnamese names without diacritics for tones throughout the thesis. Phonetic transcription of the data is provided when necessary. Note that each name in Vietnamese bears the diacritic of its tone: ngang, huyê̂n, sê̋c, nang, hoi, nga. The sacs and nangs have the same name in all environments. See Chapter Four.
(9) Minimal pairs of tone

<table>
<thead>
<tr>
<th>Phonetics</th>
<th>Tone</th>
<th>Orthography</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. lvn</td>
<td>(ngang)</td>
<td>län</td>
<td>unicorn/ neighbour/ near</td>
</tr>
<tr>
<td>2. lvn</td>
<td>(huyen)</td>
<td>län</td>
<td>turn/ time/ layer/ to totter away</td>
</tr>
<tr>
<td>3. lvn</td>
<td>(sac1)</td>
<td>län</td>
<td>to surpass/ to jostle</td>
</tr>
<tr>
<td>4. lvn</td>
<td>(nang1)</td>
<td>län</td>
<td>to tuck/ to defraud</td>
</tr>
<tr>
<td>5. lvn</td>
<td>(hoi)</td>
<td>län</td>
<td>to slip away, to escape</td>
</tr>
<tr>
<td>6. lvn</td>
<td>(nga)</td>
<td>län</td>
<td>crack-brained/ to confuse</td>
</tr>
<tr>
<td>7. lvt</td>
<td>(sac2)</td>
<td>lät</td>
<td>unstable, unreliable (lvt lvt)</td>
</tr>
<tr>
<td>8. lvt</td>
<td>(nang2)</td>
<td>lät</td>
<td>to overturn/ to capsize/ chestnut</td>
</tr>
</tbody>
</table>

In this chapter I focus only on the first six tones since they occur in the same environments, i.e., ngang, huyen, sac1, nang1, hoi and nga.

Figure 1 shows pitch tracks of the six tones in the open syllable [ta] from Nguyen and Edmondson 1997. In this figure, the transcribed vowel carries a diacritic for tone.3

Figure 1. Six tones in an open syllable [ta] (Nguyen and Edmondson 1997:6)

Fig.1a. ngang and huyen
Fig. 1b. $sac_1$ and $nang_1$

![Graph showing tones $sac_1$ and $nang_1$.]

Fig. 1c. $hoi$ and $nga$

![Graph showing tones $hoi$ and $nga$.]

Tones are generally presented in pairs for both phonetic and phonological reasons. The first pair of tones, in Figure 1a, are level tones, $ngang$ and $huyen$. $Ngang$ is high in pitch and $huyen$ is low. The second pair of tones, in Figure 1b, are $sac_1$ and $nang_1$. $Sac_1$ is a high rising tone, $nang_1$ is a low falling tone ending with a glottal stop. The last pair of tones, in Figure 1c, are $nga$ and $hoi$. $Nga$ is a high falling-rising tone, broken by a glottal stop in the middle of the tone. $Hoi$ is a low falling-rising tone.

These phonetic descriptions yield two parameters which are generally used to

---

3 It is not clear from Nguyen and Edmondson whether the values given in Figure 1 are averaged across speakers or are from one speaker.
describe Vietnamese tones (Doan 1977, Hoang Cao Cuong 1989 among others). The first refers to tonal shape: level \(ngang, huyen\), contour \(sac1, nang1\) or curve \(hoi, nga\). The second parameter involves tonal height: high \(ngang, sac, nga\) and low \(huyen, nang, hoi\). These tones are thus generally classified as follows.

(10) Traditional classification of tones

<table>
<thead>
<tr>
<th></th>
<th>level</th>
<th>contour</th>
<th>curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>(ngang)</td>
<td>(sac1)</td>
<td>(nga)</td>
</tr>
<tr>
<td>low</td>
<td>(huyen)</td>
<td>(nang1)</td>
<td>(hoi)</td>
</tr>
</tbody>
</table>

2.2. Proposal: a structural representation

In this section I investigate the representations of Vietnamese tones based on synchronic evidence. I first summarize how the two terms 'contour' and 'register' are used in the literature as tonal features, then I discuss how the terms are used in this study before presenting the proposed model.

2.2.1. Tonal features: 'contour' and 'register'

In the literature on tone, including standard works on phonetics, the terms 'contour' and 'register' are not very well-defined. Generally, 'contour' refers to tonal shape while the definition of 'register' is largely based on tonal height.

Pike 1948 divides tone languages into two categories, register tone languages and contour tone languages, depending on which feature of pitch behaviour is significant in the language. Laver (1994:465) proposes three types of tone languages. The first type is the register tone system. This is a system where the relevant feature of word-identifying
pitch behaviour is the relative height of the syllabic pitches within the speaker's pitch-span. The second type is the contour tone system. This is a system where the relevant feature is less the relative height of the tone, and more *its shape* or trajectory, together with its general placement in the speaker's pitch-span. The last type is a mixed register/contour tone system. This is a system where the end point cannot be identified directly with any of the level tones. With this definition this type is not a concern here.

In the literature on tones, representations of tones and tonal features almost invariably reflect the assumption that register is pitch. Yip 1980 proposes two features for tone, as in (11): a feature [Upper] for register, and a feature [High] for tone. The feature [Upper] is a binary feature, which, according to Snider (1999:152), is either high or low, indicating the tone is in the higher or lower register.

(11) 

<table>
<thead>
<tr>
<th>Register: Upper</th>
<th>Raised H</th>
<th>H</th>
<th>Mid</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch: High</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

(Yip 1980)

Bao (1999b:3) states that register is the pitch level of a tone and contour is how the pitch behaves over the duration of the tone-bearing unit; that is, register is static and the contour is dynamic. With the assumption that register is tonal pitch, Bao (1999a:487) describes a phenomenon called 'Register harmony' in tone sandhi in Chaozhou (Chinese) as follows: 'the pitch height of the sandhi tone - its *register* - is determined by the register of the following tone'.
The assumption that register is tonal pitch is also shared in the work on African tone languages of Snider 1999 and others. Clements 1981 proposed the representations of tones in (12). The register feature in the first row indicates that the tone is in the upper or lower register and the feature in the second row indicates the higher or lower tone within that register.

(Clement 1981)

<table>
<thead>
<tr>
<th>Register:</th>
<th>Raised High</th>
<th>High</th>
<th>Mid</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>h</td>
<td>h</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td>Pitch:</td>
<td>h</td>
<td>l</td>
<td>h</td>
<td>l</td>
</tr>
</tbody>
</table>

In short, both concepts ‘register feature’ and ‘tone feature’ seem to include tonal height. The ‘register feature’ indicates a pitch range, usually divided into two registers, high and low. The ‘tone feature’ refers to the pitch of that tone within its register, e.g., a high tone in a low register is a low tone that is articulated in the highest point within the low register.

I will adopt the assumption that ‘contour’ indicates tonal shape. This is also a common assumption in the Vietnamese literature, as we will see in section 2.4.1. However, I propose a definition of ‘register’ that does not involve pitch height. I use ‘register’ to represent different phonation types (see Chapter Five). Although this use is non-standard, I will use the term since it is familiar. In Chapters Three and Five, we will see that the register feature is, in fact, a laryngeal feature.
2.2.2. Representation of Vietnamese tones

In this section, I present a model of Vietnamese tones in which tones are grouped together according to their markedness relationships, and I provide phonological evidence for the structure proposed. The phonetic motivation for this model is presented in Chapters Three and Five.

I assume that tonal features belong to a level of structure which is independent of segmental features (e.g., Leben 1973, Goldsmith 1976, Odden 1995). I assume two major features in tonal representation, namely Register and Contour. I assume here that Register is not tonal pitch, but indicates phonation type. Specifically, it refers to modal voice, creaky voice and breathy voice. Because this assumption is so distinct from the common assumption that register is pitch, the concept ‘register’ in this study must be tested phonetically (see Chapters Three and Five).

In this study, the term ‘contour’ refers to tonal shape, indicating whether a tone is flat (level tones), or whether there is some movement during the course of the tone. Movement can be of two types. It can involve a change of direction once, e.g., going up or down to make a rising or falling tone, or a change of direction more than once, e.g., going down and then up to make a falling-rising tone. Tonal contour, therefore, does not refer to tonal height, e.g., a ‘level’ contour might be a ‘high’ or ‘low’ tone, or a ‘falling’ tone might be ‘high’ in the pitch range. It instead refers simply to shape.

I argue for a tonal representation in Vietnamese in which the structural relationship between Register and Contour is one of sisterhood, not dominance (see section 2.3.1 for arguments). The model that I propose is shown in (13). Contour and
Register are two organizing nodes, as anticipated by the discussion in Chapter One, dominated by a tone node.

(13)  

\[ T \]

Contour Register

This model is similar to the one proposed in Bao 1999b, but I will develop it further in this study.

The model in (13) predicts that some phonological processes can affect the whole tone (T) while others might affect only Register but not Contour, and vice versa. In the following sections, I use neutralization as evidence for the independence of Contour and Register features, and for constituency in the feature geometry of tones. In particular, I argue that tones are organized hierarchically. I propose the more highly articulated structure of Vietnamese tones given in (14). For the moment, the terms used here are not well-defined, but in Chapter Five these terms will be discussed and justified. Some terms in (14) are for descriptive and classificatory purposes, e.g., [non-level], [curve]. However, the formal representation of each tone will immediately follow in (18). The tonal root node in (13) is represented by the tone's name in (14), and later in (18), for convenience.
(14) Structure of Vietnamese tones

a. \textit{ngang} & \textit{huyen} \\
   | & R \\
   | & [laryngeal] \\
   | & [spread] \\

b. \textit{sac1} & \textit{nang1} \\
   | & C \\
   | & [non-even] \\
   | & [non-even] & [laryngeal] \\
   | & [constricted] \\

c. \textit{hoi} & \textit{nga} \\
   | & C \\
   | & [curve] \\
   | & [curve] & [laryngeal] \\
   | & [constricted] \\

21
Before developing this model in greater detail, I lay out a set of assumptions that I adopt concerning representations. In (14), the two major components of tones are Contour (C) and Register (R), as in (13) above. I use the term [laryngeal] as a marked feature for register, which appears in huyen, nang1 and nga. The feature [laryngeal] has [spread] as its dependent in huyen, and [constricted] as its dependent in nang1 and nga. The evidence for the register features and their phonetic realization will be discussed in Chapter Five. With respect to the contour feature, in (14a), the tones ngang and huyen are essentially flat. They are not specified as [even] for the contour feature. The feature [non-even] (changing direction) appears in (14b) for the rising and falling tones, sac1 and nang1. The feature [curve] appears in (14c) for two curved tones (i.e., it goes down and then up), hoi and nga.

Before I present the formal structure of tones in (18), let me explain how the contour feature is specified. I will assume that tones can be captured through the use of what I call points, notated by •. These points basically indicate whether there is movement during the tone. A single point indicates a level tone. Because 'contour' shows the F0 value of a tone during the time of production, I assume that at least one point needs to be specified for every tone to generate the tone contour. If the tone is level, i.e., there is no movement during its time course, one point is the beginning point (onset), marked with •, and it generates the end point (offset) as illustrated in (15). The register feature is omitted here for convenience. The underlying contour feature is specified for one point in (15a). While a level tone has only a single point phonologically, it phonetically nevertheless extends over a tone period. On the surface, (15b), this point generates the second point ending the tone.
I also assume that if there is movement during the tone, two points at the phonological level are needed to indicate the contour (movement) of the tone, e.g., whether rising or falling. I assume, therefore, two points at the phonological level must be interpreted as having different values. The movement from the onset to the offset of a rising or falling tone is illustrated in (16). Note that I do not use the height of the points in any formal sense here in the surface forms, but only simply for illustrative purposes.

In (16a), a contour tone (a tone with movement) has two points phonologically. If it is a rising tone, phonetically the second point is higher than the first point to indicate the rise (16b). If it is the falling tone, the second point is lower than the first point on the surface to represent the fall.
If the tone is a curved tone, e.g., the tone goes down and rises up, three points are needed to represent the complex movement phonologically, as illustrated in (17a).

(17) Specification for contour of a tone that changes direction more than once

a. Underlying

\[
\begin{align*}
\text{C} & \\
\rightarrow & \\
\cdots & \\
\end{align*}
\]

b. Surface

\[
\begin{align*}
\text{C (concave)} & \\
\rightarrow & \\
\cdots & \\
\end{align*}
\]

A concave tone is phonetically represented with the mid point lower than the first and last points to show the tone goes down and then up (17b). A convex tone is represented phonetically with the mid point higher than the first and last points to show the tone goes up then down (17c). The phonetic interpretation of contour and register features will be discussed further in Chapter Five, where I argue that pitch is derived from the register feature and the contour feature.

Based on (14) to (17), the phonological representation of each tone is shown in (18). I will show in Chapter Five, section 5.2.2. how the particular directions of the contours are derived.
(18) Formal structure of Vietnamese tones

a. \( \text{ngang} \)
   \[
   \begin{array}{c}
   & \bullet \\
   & \mid \\
   C & \mid \\
   \end{array}
   \]

b. \( \text{sac1} \)
   \[
   \begin{array}{c}
   & \bullet \\
   & \mid \\
   C & \mid \\
   \end{array}
   \]

c. \( \text{hoi} \)
   \[
   \begin{array}{c}
   & \bullet \\
   & \mid \\
   C & \mid \\
   \end{array}
   \]
Assuming Contrastive Underspecification (Avery & Rice 1989, Rice 1992, Wu 1994 and others), i.e., an unmarked feature is absent underlingly unless there is a contrast which forces it to be present, following this assumption, unmarked features are not present in (18). For instance, in terms of register, the feature [laryngeal] is marked; therefore, it is present underlingly in huyen but absent in ngang (18a).

In (18a) the level tone ngang does not have any specified features. Its counterpart, huyen, has the register feature [laryngeal] with [spread] (breathy, see Chapter Five) as its dependent. In each level tone, one point is specified for the contour feature, marked with (•), as shown in (17). This point does not need to be specified as ‘h’ or ‘l’, because this value is predictable from the laryngeal feature. See the discussion of the phonetic realization of tone in Chapter Five.

In (18b), sac1 does not have a register feature but it has a contour. As there is a movement in this tone, there are two points to represent the contour: two specified points marked with (•). Its counterpart, the falling tone nang1, has the register feature [laryngeal] with the feature [constricted] (creaky, see Chapter Five) as its dependent. On the contour side, nang1 is specified for two points to show the movement. (18c) shows the representation of the two curved tones, hoi and nga. In (18c) hoi does not have the register feature. It has a contour, specified for three points to generate a curved tone. Its counterpart, nga, has features on both sides. On the register side nga has a laryngeal feature with [constricted] as its dependent. On the contour side, as in hoi, nga has three points specified to represent the curve.

The representation in (18) shows that each pair of tones shares the same contour: ngang and huyen only have one point specified phonologically; sac1 and nang1 have two
points specified phonologically, showing that the tone changes in direction; *hoi* and *nga* both have three points specified phonologically to show that the tone changes its direction more than once, i.e., it goes down and then rises (there are no rising-falling (lhl) contours in Vietnamese. On the register side, the representation in (18) shows that tones on the left, *ngang*, *sac1* and *hoi*, do not have the register feature. Tones on the right share the same register feature, [laryngeal] with a dependent.

Markedness relations of tones are based on structural complexity, as discussed in Chapter One. Four kinds of markedness apply to tones: markedness within register, markedness within contour, markedness between pairs and markedness within pairs. The representations in the left column of (18) show that the unmarked register consists of tones that are unspecified for register, i.e., *ngang*, *sac1* and *hoi*. In the right column, the marked register consists of tones that are specified for register with the feature [spread] or [constricted], i.e., *huyen*, *nang1* and *nga*. Tones in the right column, therefore, are more complex than those in the left column with respect to register, and thus they are more marked according to (7) in Chapter One, repeated here in (19).

(19) Horizontal complexity

<table>
<thead>
<tr>
<th></th>
<th>complex</th>
<th>simple</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>E</td>
<td>D</td>
</tr>
<tr>
<td>F</td>
<td>G</td>
<td>F</td>
</tr>
</tbody>
</table>

27
For instance, *ngang* and *huyen* have an equal degree of complexity on the contour side, with only one specified point for each; however, on the register side *huyen* has the feature [laryngeal] with [spread] as its dependent. In this case, *huyen* is horizontally more complex than *ngang*.

In terms of contour, between pairs, *ngang* and *huyen* are simpler than *sac1* and *nang1*, which in turn are simpler than *hoi* and *nga*. Let us compare *ngang* and *sac1* first. Neither of these two tones has features under the register node. However, on the contour side, *sac1* has two points specified for the contour feature while *ngang* only has one point. Therefore, *sac1* is horizontally more complex than *ngang* (with two points of contour). Similarly, between *huyen* and *nang1*, *nang1* is vertically more complex than *huyen* (with two points specified for the contour). On the register side, although both *huyen* and *nang1* have the [laryngeal] feature, I assume that [constricted] is more marked than [spread], as discussed in Chapter Five. Note that this is a different type of markedness, based on substance rather than structure. Next we look at the pairs *sac1* and *nang1* with *hoi* and *nga*. On the register side, *sac1* and *hoi* are equally complex: they do not have the register feature. *Nang1* and *nga* are equally complex: they both have the feature [laryngeal] with [constricted] as its dependent. However, on the contour side, *sac1* and *nang1* are horizontally less complex than *hoi* and *nga*, respectively. *Hoi* and *nga* have three points specified for the contour feature compared to two in *sac1* and *nang1*.

Overall, the simplest structure is *ngang*, having only one contour point specified. The most complex tone is *nga* with the laryngeal feature [constricted] and three points specified for the contour node.

The markedness relation of tones between registers is summarized in (20).
(20) Markedness of tonal registers based on complexity

<table>
<thead>
<tr>
<th>Unmarked register</th>
<th>ngang</th>
<th>sac1</th>
<th>hoi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marked register</td>
<td>huyen</td>
<td>nang1</td>
<td>nga</td>
</tr>
</tbody>
</table>

Within each pair in (20), the tone on the upper row featurally is less complex than its lower counterpart according to the complexity measure in (19), with the complexity resting on the addition of the register feature. For instance, in (18a), ngang does not have the register feature while huyen has the feature [spread] under Register. The markedness relations of tones within their pair in the opposite registers is summarized in (21), where ‘<’ indicates that the lefthand member is less complex or less marked than the righthand member.

(21) Markedness of tones within each pair

\[ \text{ngang} < \text{huyen}, \quad \text{sac1} < \text{nang1}, \quad \text{hoi} < \text{nga} \]

In terms of contour, markedness of tones is also evaluated among tonal pairs across registers. Of the three pairs, ngang-huyen is the least marked (18a). These tones have only one point specified for the Contour node. The second least complex pair is sac1 - nang1 (18b). Both have two points specified for the contour node and sac1 unspecified for register. This pair is horizontally more complex than the pair ngang-huyen according to (19). Finally, the most complex pair hoi-nga (18c) is horizontally more complex than sac1 - nang1 with three points specified for the contour node. The markedness relations among tone pairs are summarized in (22).
Markedness of tone pairs the least to most marked

\[ ngang - huyen < sac1 - nang1 < hoi - nga \]

2.2.3. Summary

I have proposed that Vietnamese tones are grouped into two registers, an
unmarked register consisting of \textit{ngang}, \textit{sac1} and \textit{hoi} and a marked register consisting of
\textit{huyen}, \textit{nang1} and \textit{nga}. Across registers, tones are grouped in pairs according to their
contour: \textit{ngang}-\textit{huyen}, \textit{sac1}-\textit{nang1} and \textit{hoi}-\textit{nga}. In terms of complexity, tone is marked in
different dimensions: (i) within each pair, one tone is horizontally more complex than
the other; (ii) across pairs, one pair is horizontally more complex than the other. The pair
\textit{hoi}-\textit{nga} is the most marked, and \textit{ngang}-\textit{huyen} is the least marked.

2.3. Evidence for constituents and markedness relations of tones

Diagnostics for markedness are often discussed in the phonological literature.
Two that are often mentioned are frequency and neutralization targets. With respect to
frequency, a less marked feature often has wider distribution than a more marked one
With respect to neutralization targets, there is avoidance of marked features: unmarked
features are retained in neutralization environments and marked features are lost,
merging with unmarked ones.

In this section, I present evidence for the structure of tones proposed in (18) and
the markedness relations of tones shown in (20) to (22). The evidence is drawn mainly
from tone frequency, neutralization in dialects, reduplication, and poetry. Support for
the proposal also comes from clitics and borrowings. The markedness relationship between tonal pairs predicts that when neutralization occurs, it will tend to neutralize the more marked tones to the less marked tones. With respect to frequency of occurrence, the markedness relationship predicts that the least marked tones will tend to occur more often than the more marked tones.

2.3.1. Reduplication and tone harmony: evidence for the independence of register and contour

In the structure in (7), Contour and Register have a sisterhood relation. This model predicts that there might be phonological processes which affect only Contour, leaving Register intact, and vice versa. Unlike many tone languages, Vietnamese has no tonal processes such as assimilation or tone sandhi. However, an investigation of reduplication yields insight into the phonology of tones. Vietnamese reduplication consists of several patterns including total and partial reduplication in both productive and unproductive processes. I will not attempt to provide a detailed description of segmental processes, but focus only on tonal patterning (for details of reduplication processes, see Burton 1992, Huu & Vuong 1980, Truong 1970, Ngo 1984, Thompson 1965, Vuong & Hoang 1994).

In reduplication, tones pattern in a predictable way, providing evidence for the two registers, namely unmarked (traditional high) and marked (traditional low). The data in (23) show the pattern of unmarked register tones and those in (24) show the pattern of marked register tones in a productive process of reduplication. The base tone is underlined. The tone of the second syllable is not predictable, but its register is. For
instance, in (23a), the base tone is *ngang*, an unmarked tone; the reduplicant must be either *sac1* or *hoi*. In this case, it is *sac1*. In (23), the base tones all are unmarked register tones, i.e., *ngang, sac1* and *hoi*; the reduplicant tone must also be from the unmarked register, i.e., *sac1* in (23a), *hoi* in (23b), and *ngang* in (23c).

(23) UNMARKED register: *ngang* - *sac1* - *hoi*

a. *mau (ngang)* ‘fast’ > *mau man (ngang - sac1)* ‘very fast’

b. *lau (sac1)* ‘clever’ > *lau linh (sac1 - hoi)* ‘very clever’

c. *do (hoi)* ‘red’ > *do dan (hoi - ngang)* ‘very red’

Thus, in this type of reduplication, the contour of the reduplicant tone (second syllable in (23)) differs in an unpredictable way from that of the base. However, the register of the base is always replicated in the reduplicant.

In (24), the base tones are marked register tones, i.e., *huyen, nang1* and *nga*, and the reduplicant tone must be a marked register one, i.e., *nang1* in (24a), *nga* in (24b), *huyen* in (24c). Tone harmony is never violated.4

(24) MARKED register: *huyen* - *nang1* - *nga*

a. *tan (huyen)* ‘worn out’ > *tan ta (huyen - nang1)* ‘worn out’

b. *lanh (nang1)* ‘cold’ > *lanh leo (nang1 - nga)* ‘very cold’

c. *mo (nga)* ‘grease’ > *mo mang (nga - huyen)* ‘very greasy’

---

4 These forms are lexicalized and the process is not productive. However, register preservation is also strictly kept.
Register preservation in tone harmony is strictly obeyed even when there is no distinction between certain tones in some dialects. For instance, in Southern dialects nga is neutralized to hoi (see Section 2.3.5.1 below). Because the distinction between hoi and nga is still expressed in the orthography and in productive reduplication (see 2.3.2 below), in order to decide whether a morpheme has the tone hoi or nga, Southerners have to memorize the rule of ‘ngang-sac-hoi’ and ‘huyen-nang-nga’ (Hoang Thi Chau 1989:201). For example, in ‘nong nay’ (temper), if ‘nong’ has sac1, ‘nay’ must be hoi, which is in the same register as sac1. It would be interesting to test if Southern speakers still maintain register with nonsense words, and if there might be some agreement on the contour of the reduplicant.

To summarize, in reduplication register must be retained in tone harmony, although contour can vary. Tones in the unmarked register pattern together in a group (ngang, sac1, hoi), and tones in the marked register form a group as well (huyen, nang1, nga). This patterning provides evidence for two registers in Vietnamese. These registers are independent of contour features as tonal patterning in reduplication only affects Contour but not Register, e.g., an unmarked register tone can pattern with another unmarked tone which has a different contour feature.

In the tonal literature, other researchers including Bao 1999a,b, Duanmu 1990 and Snider 1999 also argue for Register and Contour being sisters. Yip (1989, 1995) proposes a model of tonal representation in which a register feature dominates a pitch feature, arguing that tones spread as a whole. Below is a high rising tone in Yip’s model. ‘H’ stands for register; ‘l, h’ stand for pitch.
Register dominates Pitch (Yip 1989)

\[
\begin{array}{c}
\sigma \\
| \\
H \\
\hline
l \quad h
\end{array}
\]

The fact that Register is preserved in Vietnamese reduplication is a challenge for this organization, regardless of the formal details of how tones are determined in reduplication, either by spreading or copying of register. If Register dominates Pitch as in (25), the reduplicant tone would have the same pitch feature (i.e., 'Contour' in the present work) as its base, but this is not the case in Vietnamese, where reduplicant tones do not share contour features with their base tones.

Yip (1995:487) also mentions that spreading in African tone languages usually shows that voicing and Register do not interact, where high and low tones spread freely across syllables, across any kind of consonant (for instance, Luganda from Hyman, Katamba and Walusimbi 1987). This type of spreading shows that only Pitch is involved, not Register. From this fact, Yip suggests that there are two different types of dependency relations between Pitch and Register in Chinese and African languages.

With the model in (13), where register and contour are independent, it is possible to suggest a straightforward account for both types of spreading: in African languages only Contour spreads; in Vietnamese reduplication, only Register spreads.
2.3.2. Reduplication: evidence for the markedness of ngang and huyen and for register being independent of contour

Reduplication, as discussed in 2.3.1, provides evidence for the independence of contour and register features. In addition, it gives evidence for the tonal pair ngang-huyen being unmarked compared to the other tones.

It is frequently discussed in the literature that in certain positions, only a limited number of segments or features is allowed (see Rice 1996, Paradis and Prunet 1989 among others). For instance, almost all segments in the inventory of a language are allowed in syllable-initial position, but fewer tend to occur in syllable-final position. In Korean, all coronal obstruents, regardless of their particular coronal place and their manner features, are neutralized to [t] in coda position (Cho 1991:171). In the Saigon dialect of Vietnamese, either coronals or velars can surface syllable-finally while all consonants occur initially, (except /p/ (Pham 1998)). With respect to laryngeal features, German allows only voiceless obstruents in final position; voiced obstruents undergo devoicing in this position (e.g., Brockhaus 1995, Jessen 1998). Features that are required in certain environments are often argued to be unmarked. Therefore, in the examples involving coronal consonants, where [coronal] is the only place allowed in the coda, [coronal] is argued to be unmarked. See Greenberg 1966 and others for discussion of neutralization as a characteristic of lack of markedness.

Assuming that unmarked features occur in neutralization environments, the structure in (18) predicts that if there are tones occurring in such an environment, they should be ngang or huyen, the least marked ones. And indeed, this is the case - ngang and
huyen are found in neutralizing environments. I look at this first with respect to productive reduplication.

There is another process of reduplication in Vietnamese, this time a productive one, in which the base tone is neutralized to ngang or huyen. The particular form depends on the register of the base tone. If the base tone is an unmarked register tone (sac1 or hoi), the reduplicant tone must be ngang, another unmarked register tone (26a). If the base tone is a marked register tone (nang1 or nga), the reduplicant tone must be huyen, another marked register tone (26b). Recall that the base is underlined.

(26) Reduplication with the reduplicant tone ngang or huyen

a. Unmarked register base, ngang in reduplicant

\[
\begin{align*}
\text{trang} & (\text{sac1}) \quad \text{‘white’} & > & \text{trang trang} & (\text{ngang - sac1}) \quad \text{‘rather white’} \\
\text{do} & (\text{hoi}) \quad \text{‘red’} & > & \text{do do} & (\text{ngang - hoi}) \quad \text{‘rather red’}
\end{align*}
\]

b. Marked register base, huyen in reduplicant

\[
\begin{align*}
\text{lanh} & (\text{nang1}) \quad \text{‘cold’} & > & \text{lanh lanh} & (\text{huyen - nang1}) \quad \text{‘rather cold’} \\
\text{se} & (\text{nga}) \quad \text{‘soft’ (voice)} & > & \text{se se} & (\text{huyen - nga}) \quad \text{‘rather soft’}
\end{align*}
\]

The proposed model offers an account for this process as follows. The realization of the

5 There is an alternation if the base tone is nga. The reduplicant can be either huyen or ngang (Thompson 1965, Ngo Thanh Nhan 1984, Hoang Dung 1998). It is sometimes argued that this exception can be explained by the fact that hoi originally was in the low register and changed to the high register (see Section 3.2.2); therefore, huyen became ngang (Nguyen Tai Can 1977:118). This analysis is not a satisfactory one, as it does not explain why the other option does not occur: there is no alternation if the base tone is hoi (25a), the reduplicant is always ngang, not huyen.
reduplicant tone is shown in (27). In this process, the reduplicant tone is pre-specified as a level tone - its contour is specified. Its register is predictable, being the same as the register of the base (there is no register feature in (27a), but a marked register in (27b)).

(27) Processes of neutralization in reduplication
a.  trang (sac1) 'white' > trang trang (ngang - sac1)

b.  lanh (nang1) 'cold' > lanh lanh (huyen - nang1)

Details are discussed in Chapter Five; the important issue here is that in neutralization environments, less complex contours are found. The less complex contour than rising and falling is level. This process shows that ngang and huyen are two unmarked tones, as they occur in a neutralizing environment.
2.3.3. Frequency of occurrence: evidence for the markedness of tones

Frequency of occurrence can be a diagnostic for the markedness of tones. I assume that a less marked feature occurs more frequently than a more marked one (Greenberg 1966, Maddieson 1984, Hamilton 1996 and others). The ranking of markedness of tones in (20), (21) and (22) predicts that among tones, the least marked tone, n gang, should have the highest frequency and the most marked one, nga, should have the lowest frequency. Among pairs, the least marked pair n gang-huyen occurs the most and the most marked one hoi-nga occurs the least.

Vo Xuan Hao (1997:20) examined 4243 monosyllabic words in a dictionary. Leaving out syllables ending in stops -p, -t, -k, he provides the raw numbers of occurrence for each tone in sonorant-final syllables as follows.

(28) Frequency of tones in Vo (1997) excluding -p, -t, -k

<table>
<thead>
<tr>
<th>Total</th>
<th>n gang</th>
<th>huyen</th>
<th>sac1</th>
<th>n ang1</th>
<th>hoi</th>
<th>nga</th>
</tr>
</thead>
<tbody>
<tr>
<td>4243</td>
<td>1029</td>
<td>840</td>
<td>845</td>
<td>606</td>
<td>570</td>
<td>353</td>
</tr>
</tbody>
</table>

The table in (28) shows a striking match between frequency and markedness. The unmarked tone n gang occurs more than its more marked counterpart huyen (1029 and 840 respectively). Similarly, sac1 occurs more than its more marked counterpart n ang1 (845 and 606 respectively). Finally, hoi occurs more than its more marked counterpart nga (570 and 353 respectively). The ranking of markedness within pairs given in (21) is repeated here as (29).
Assuming that frequency of occurrence is a diagnostic for the markedness of tones, the ranking in (29) predicts that the least marked tone, \( \text{ngang} \), has the highest frequency and the most marked one, \( \text{nga} \), has the lowest frequency. Among pairs, the least marked, \( \text{ngang-huyen} \), occurs the most often and the most marked one, \( \text{hoi-nga} \), occurs the least. These predictions about frequency in Vietnamese are borne out, providing evidence for the markedness relations proposed here.

2.3.4. Poetry: more evidence for \( \text{ngang} \) and \( \text{huyen} \) sharing the same contour features

Poetry provides another piece of evidence that is compatible with the claim that \( \text{ngang} \) and \( \text{huyen} \) forming a tonal pair. In a traditional verse called six-eight verse, there is a pattern of alternation between even tones, i.e. Vietnamese ‘bằng’, (\( \text{ngang} \) and \( \text{huyen} \)) and the so-called sharp tones, Vietnamese ‘trắc’, (all the other tones). This verse has a six-syllable line followed by an eight-syllable line. The two lines form a pair. A poem can have any number of pairs of lines. Patterns of tones in the six-eight verse are shown in (30). ‘E’ stands for even tones (\( \text{ngang, huyen} \)), ‘N’ stands for non-even tones. The rhymed position is underlined. Tones in the first, third and fifth positions are flexible with respect to whether they house even or sharp tones. Tones in the second, fourth, sixth and eight positions must strictly follow the pattern. For instance, a tone in the first

\[(29) \quad \text{ngang} < \text{huyen}, \text{sacl} < \text{nang1, hoi} < \text{nga} \]
position can be either E or N, but a tone in the second position must be E (ngang or huyen). Thus only even tones can occur in the rhymed position, underlined in (30).\(^6\)

(30) Pattern of tones in the six-eight verse

\[
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
E & E & N & N & E & E & & \\
E & E & N & N & E & E & N & E \\
\end{array}
\]

The fact that tones in the six-eight verse pattern in two groups, one group consisting of ngang and huyen, and the other consisting of the other tones, shows that ngang and huyen form a pair. This pair shares the same contour feature (both are level tones).

---

\(^6\) In the example below, the words in rhyme position are in italics and underlined. English translation is taken from Huynh 1973.

```vietnamese
Trăm năm trong cõi ngubi *tu* (ngang)
Chữ tài chữ mềm khéo *lá* (huyen) ghẹt *hâu* (ngang)
Trải qua một cuộc bè *dầu* (ngang)
Nhưng điều trong thầy mà *dầu* (ngang) đốn *lông* (huyen)
Là gi bị sắc tử *phòng* (ngang)
Trời xanh quen tôi mà *hồng* (huyen) đánh *ghen* (ngang)...

Nguyễn Du (Truyện Kiều)
```

'A hundred years - in this life span on earth, how apt to clash, talent and destiny! Men's fortunes change even as nature shifts - the sea now rolls where mulberry fields grew, one watches things that make one sick at heart. This is the law: no gain without a loss, and Heaven hurts fair women for sheer spite...' (Huynh 1973).

Note that all the tones in rhyme positions are level, either ngang or huyen.
Two further pieces of evidence show that ngang and huyen are unmarked tones. First, in terms of frequency, these tones occur much more often than non-even tones in verse, as discussed above. Second, rhyme is found in positions where only even tones can occur: the last syllable of the six-syllable line rhymes with the sixth syllable of the eight-syllable line; the eighth syllable of this line then rhymes with the last syllable of the next six-syllable line, and so on. With the dominant occurrence of even tones, which do not have complex contours, the six-eight verse is the smoothest verse in Vietnamese poetry in terms of sound harmony. Since only even tones can occur in rhyme position, one can assume that because this position is heavy enough, i.e., metrically prominent, an unmarked tone is the best candidate to keep the general smoothness of the poem as compensation for the heavy position. This neutralization is not active since it does not change a lexical tone (an active neutralization changes a lexical tone such as some processes in reduplication), but rather it is a choice of words with tones ngang and huyen. The fact that only ngang and huyen can occur in this position and that they pattern as a group supports the claim that they are unmarked tones and form a contour pair.

---

7 In modern poetry, sometimes the alternative rule is deliberately violated, usually by great poets, to express a strong or agonizing feeling that would not be effectively expressed in the smooth six-eight verse. For instance, in the following sentences (Nguyễn Bình - Lỡ bước sang ngang), the author describes a girl’s sorrow when her lover left. In the second position of the first line, instead of an even tone, he uses sacl, a non-even tone (mặt), and in the fourth position, instead of a non-even tone, he has ngang, an even tone, (dôi).

Up mặt (sac) vào đôi (ngang) bàn tay (Burrying her face in her hands)
Chỉ tôi khúc sọt một ngày một đêm (My sister did not stop crying for the whole day and night)
Nguyễn Bình (Lỡ bước sang ngang)

However, the rhyme position is never violated, containing 'tay' (ngang) and 'ngày' (huyen).
2.3.5. Neutralization in dialects: evidence for markedness among pairs and between tones in a pair.

One possibility predicted by the model in (18) is that if neutralization takes place, the more marked member of a pair will be neutralized to the less marked member. Moreover, because Contour and Register are separate from each other, the model also predicts the following types of neutralization: in one type, Contour contrasts are retained but Register is neutralized; in another type, Contour contrasts are lost but Register distinctions remain. These various patterns of neutralization are found in Southern and many Central dialects of Vietnamese. Note that this section shows the phonetic representation of tones in neutralization processes. See the discussion of how tones are realized in Chapter Five.

2.3.5.1. Neutralization within Contour- hoi and nga: markedness evidence.

One type of neutralization, where Contour remains but Register is neutralized, is found in Southern and many Central dialects. In terms of contour, the markedness relation between the pairs in (20) predicts that neutralization might occur with the most marked pair, hoi and nga. In terms of register, the markedness relation between the tones of each pair in (21) predicts that, because hoi is less marked than nga, hoi would be retained in neutralization and nga lost, merging with the unmarked hoi. Indeed, this case
is found in Southern and Central dialects. In these dialects, *hoi* and *nga* neutralize, resulting in *hoi* (Doan 1977, Ngo 1984 and others).\(^8\)

The hierarchical analysis of tones proposed in section 2.2.2 offers a straightforward and simple account of the neutralization of *hoi* and *nga*. In this process, the contour contrast remains but the register contrast is lost (delinking of the register feature [laryngeal]). This is shown in (31).

(31) Neutralization from *nga* to *hoi*

```
nga                  \[\text{Delinking feature of R}\]
\[\text{C} \quad \text{R} \rightarrow \quad \text{C}\]
\[\text{\[laryngeal\]}\]
```

2.3.5.2. Neutralization within Register: markedness evidence between pairs

I now turn to evidence for markedness relations between tones in the same register. The markedness relations of tonal register in (20) are repeated here as (32) for convenience.

---

\(^8\) Note that the neutralization processes discussed in this section are most likely diachronic rather than synchronic in nature. More work is needed on the synchronic tonal systems of the dialects discussed in this section.
Recall that neutralization can occur in two dimensions: within Contour, as discussed in Section 2.3.5.1, or within Register. As stated in (20), the markedness between the pairs of tones in (32) increases from left to right. In the unmarked register, \( ngang \) is the least marked tone and \( hoi \) is the most marked. In the marked register, \( huyen \) is the least marked tone while \( nga \) is the most marked.

In terms of Contour, (32) predicts that if neutralization of a contour occurs, it will occur with the most marked pair, \( hoi \) and \( nga \). This type was shown in previous sections. In terms of Register, (32) predicts that in neutralization within register, the lost tone in the unmarked register will be \( hoi \), and in the marked register, \( nga \). These are the two most marked tones in their registers. These predictions are correct, as shown below. Recall that the shifts discussed here are most likely diachronic.

In all Nghe Tinh dialects spoken in North Central Vietnam, neutralization occurred in the marked register: \( nga \) neutralized to \( nang1 \) (Hoang Thi Chau 1989, Vuong and Hoang 1994). The proposed model offers a simple account for this fact. In this process, Register remained intact. \( Nang1 \) resulted from delinking of one point of the contour feature in \( nga \). A complex contour (curve in \( nga \)) was neutralized to a less complex one (\( nang1 \)). Although (33) shows delinking of the last point, it should be noticed that the deleted point could be any of the points because the result would be the

\[
\begin{array}{ccc}
\text{Unmarked register} & \text{\( ngang \)} & \text{\( sac1 \)} & \text{\( hoi \)} \\
\text{Marked register} & \text{\( huyen \)} & \text{\( nang1 \)} & \text{\( nga \)}
\end{array}
\]

(32)
same: according to (18), a tone with two points specified and a register feature is \textit{nang}1. In this process the register feature was preserved.

(33) Neutralization from \textit{nga} to \textit{nang}1

\begin{align*}
\text{\textit{nga} } & \quad \text{Delinking the last point} & \quad \text{\textit{nang}1} \\
\text{C} & \quad \text{R} & \quad \rightarrow & \quad \text{C} & \quad \text{R} \\
\text{\textbullet\textbullet\textbullet\textbullet \hspace{0.5cm} [laryngeal]} & \quad \text{\textbullet\textbullet \hspace{0.5cm} [laryngeal]} \\
\end{align*}

The fact that neutralization results in \textit{nang}1 shows that \textit{nga} is more marked than \textit{nang}1.

In the unmarked register, a neutralization process is also found. In Mai Ban, another dialect of Nghe Tinh, \textit{nga} merged with \textit{nang}1 in the marked register. Interestingly, in the unmarked register, \textit{hoi} and \textit{sac}1 also neutralized, resulting in \textit{sac}1 (Hoang Thi Chau 1989, Vuong and Hoang 1994), as shown in (34).

(34) Neutralization in the Mai Ban dialect

\begin{center}
\begin{tabular}{|c|c|c|}
\hline
\textit{ngang} & \textit{sac}1 & \leftarrow \textit{hoi} \\
\hline
\textit{huyen} & \textit{nang}1 & \leftarrow \textit{nga} \\
\hline
\end{tabular}
\end{center}

This supports the claim that \textit{hoi} is more marked than \textit{sac}1, and \textit{nga} is more marked than \textit{nang}1. In this process, like that in (33), Register remained inact (no register), and the
contour contrast was lost by the delinking of one point of the contour feature in _hoi_. As in (34), this is the neutralization of a tone with a more complex contour feature (curve) to a less complex one (non-curve).

(35) Neutralization from _hoi_ to _sac1_

```
hoi
|__________________________|  Delinking the first point  |__________________________|  sac1
C

\[ \begin{array}{c}
\text{\textbackslash
\textbackslash
}\text{\textbackslash}
\end{array} \]

Again, it is not necessary that it be the beginning point in the contour feature of _hoi_ that is delinked, because the result would be the same no matter which point is lost: a tone without the register feature and with two specified points for the contour is _sac1_, according to (18).

The fact that within Register neutralization occurs with the two tones _hoi_ and _nga_ provides further evidence for _hoi_ and _nga_ being the most marked pair. Interestingly, in all dialects (as seen above), the two tones _ngang_ and _huyen_ always remain regardless of different patterns of neutralization. This argues that _ngang_ and _huyen_ are the least marked tones.

I have shown that neutralization of tones in dialects occurs in two dimensions: across register and within register. In all cases, the surface tone in neutralization is always the less marked member of a pair. Across registers, the unmarked register tone
remains (e.g., hoi in Southern dialects). Within a register, the less marked tone remains (e.g., sac1 or nang1 in Nghe Tinh dialects). This is consistent in all patterns of neutralization (i.e., in both reduplication and poetry). These patterns strongly support the claim that markedness relations between tones do exist and that the features that make up tones are structurally organized as proposed.

2.3.6. Clitics: evidence for unmarked tones

Vietnamese clitics provide another piece of evidence for the markedness of tones. Clitics show that the basic set of tones consists of ngang, huyen and sac1. Among tonal pairs, ngang and huyen are the least marked. Sac1 is the less marked member of the second least marked pair (sac1 and nang1) in the markedness relations proposed in (21).

In the construction of phrases with clitics, a word is cliticized onto an adjacent word, either to the left or to the right. After attaching to the adjacent word, all segments of the clitic are deleted but its tone remains. If the clitic associates with an obstruent, the clitic tone is realized on a homorganic nasal. If it associates with a vowel or glide, the clitic tone is realized on a geminated vowel or glide (Pham 1997b). Some data are given in (36). Data are written in phonetic symbols. Only tones of the relevant syllables are shown.

---

9 The data were collected from a recorded conversation and an interview (see Pham 1997b). Even though most of the clitics are function words (e.g., pronouns, question words, prepositions, negative particle, demonstratives, classifiers, conjunctions) there are about 25 clitics in these conversations and they occur very frequently.
Clitics and their tones are underlined. For instance, in (36a) the clitic tone is \textit{huyen}. The word \textit{[la:m }‘how’ with tone \textit{huyen} is cliticized onto the host \textit{[x\text{\textalpha}]} ‘to pray’(\textit{sac1}), which ends in a coronal nasal. The clitic also surfaces as a coronal nasal bearing its original tone \textit{huyen}. In (36b), the clitic \textit{[n\text{\textalpha}]} ‘he’ with tone \textit{sac1} cliticizes onto the preceeding host \textit{[c\text{\textalpha}]} ‘to let’ (\textit{ngang}) which ends in a vowel. The vowel is lengthened, and the clitic’s tone is realized on the lengthened part of the vowel \textit{/\textalpha/}. In Vietnamese, the underlying velar final consonant is phonetically doubly articulated, a labio-velar, i.e., \textit{/k/} is \textit{[kp]} and \textit{/\texteta/} is \textit{[\texteta m]}, after back round vowels \textit{/u, o, \textalpha/}, a process seen in (36c). In (36c), the host \textit{[x\text{\textalpha}kp]}
‘to cry’ bearing tone sac2, ends with a labio-velar obstruent [kp]. The clitic [on] (ngang) ‘inside’ cliticizes onto the final consonant [kp] of the host, assimilating to the place of the segment [p] of the labial consonant. The clitic surfaces as a labial nasal which bears tone ngang. In (36d) the clitic [kum] ‘also’ with tone nga is cliticized onto the host [an] ‘eat’ with tone ngang. The clitic surfaces as a coronal nasal bearing its original tone nga. Although sac2 and nang2 are not discussed until Chapter Four, I also give an example here in (36e). There is a single clitic bearing tone nang2: the determiner [mot] ‘a’. In (36e), the word [mot] cliticizes onto the host [ka] ‘have’ (sacl), which ends in a vowel. The vowel in [ka] is lengthened, and the clitic’s tone nang2 is realized on the lengthened part of the vowel /a/. Interestingly, because sac2 and nang2 occur only in stop-final syllables, the nang2 of the clitic [mot], realized on a vowel, has to surface as huyen. 10

Assuming that cliticization restricts the range of tones, one might predict that the most frequent tones in clitics should be the less marked tones. In fact, that is the case.

In terms of frequency of occurrence, ngang, huyen and sac1 occur most frequently in clitics. Of the 24 clitics found in the recorded conversation and interview, there is only one clitic with the tone nga (see (36d)), and only one clitic, [mot] ‘determiner’, bears the tone nang2, which surfaces as huyen (see (36e)). The rest are ngang, huyen and sac1. Of the 22 clitics, seven have ngang, six have huyen and nine have sac1.

10 The clitic [mot] in (36e) deserves some comment. The fact that nang2 surfaces as huyen in this case shows that huyen is the most similar tone to nang2; therefore it is chosen to replace nang2. In fact, huyen and nang2 are in the same register, a marked register. In addition, they both have breathiness, as will be seen in Chapter Three. As a determiner, [mot] occurs most frequently in cliticization. In the environment of cliticization, nang2, the most marked tone in the marked register (see Chapter Four), is neutralized to huyen, the least marked tone in the same register. This fact shows that the most marked tones do not occur in clitics, and that huyen is an unmarked tone while nang2 is subject to neutralization.
The clitics do not constitute a large sample due to the fact that only function (grammatical) words can be clitics. Nonetheless, we still see that the similar distribution of ngang, huyen and sac1 in clitics can be explained if we assume the structure of tones as in (18): ngang, huyen and sac1 are the least marked tones. This is consistent with the finding that the smaller the inventory, the simpler the segments it has (e.g., Maddieson 1984, Rice & Avery 1993).

2.3.7. Borrowings

Borrowings from non-tonal languages such as French or English are also compatible with the claim about markedness of tones. The proposed feature structure predicts that the tones in borrowed words should be ngang or huyen, the least marked pair. This prediction is borne out. More interestingly, most borrowings bear the tone ngang, the least marked tone. For instance, among the examples given in (37), taken from dictionaries (Nguyen 1998, Uy ban KHXH Vietnam 1988), very few forms occur with huyen, e.g., ‘cà’ in cà phê ‘coffee’, ‘mùi’ in mủi-xoa ‘handkerchief’, ‘xi’ in xi-nách-ba ‘snack-bar’ or xi-căng-dan ‘scandal’. The majority of examples have ngang. These words are commonly used in the language. Note that every syllable in Vietnamese has to have a tone. In (37), tones are given beside the orthography. Note also that sac2, the less complex tone of the pair occurs in stop-final syllables.

(37) Borrowings with ngang or huyen:

cà phê       huyen-ngang       (Fr. café)       ‘coffee’

50
<p>| mủi-xoa | huyen-ngang    | (Fr. mouchoir) | ‘handkerchief’ |
| xi-càng-dan | huyen-ngang-ngang | (Fr. scandale) | ‘scandal’ |
| xi-nách-ba | huyen-sac2-ngang | (Eng.) | ‘snack-bar’ |
| ra-di-ô | ngang-ngang-ngang | (Fr. radio) | ‘radio’ |
| po-luya | ngang-ngang | (Fr. pelure) | ‘thin paper’ |
| ca-bin | ngang-ngang | (Fr. cabine) | ‘cabin’ |
| xảng-dan | ngang-ngang | (Fr. sandale) | ‘sandals’ |
| ba-lô | ngang-ngang | (Fr. ballot) | ‘backpack’ |
| ban-công | ngang-ngang | (Fr. balcon) | ‘balcony’ |
| tác-xi | sac2-ngang | (Fr. taxi) | ‘taxi’ |
| ni-lông | ngang-ngang | (Fr. nylon) | ‘nylon’ |
| ba-tê | ngang-ngang | (Fr. paté) | ‘pate’ |
| sc-mi | ngang-ngang | (Fr. chemise) | ‘shirt’ |
| xi-măng | ngang-ngang | (Fr. ciment) | ‘ciment’ |
| cao-su | ngang-ngang | (Fr. caoutchouc) | ‘plastic’ |
| ăng-ten | ngang-ngang | (Fr. antenne) | ‘antenna/ spy (slang)’ |
| bâng | ngang | (Fr. banque) | ‘bank’ |
| banh | ngang | (Fr. panne) | ‘broke down (car)’ |
| boong | ngang | (Fr. pont) | ‘deck’ |
| bi-da | ngang-ngang | (Fr. billard) | ‘billiards’ |</p>
<table>
<thead>
<tr>
<th>Word</th>
<th>Tone</th>
<th>Original Language</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ca</td>
<td>ngang</td>
<td>(Fr. cas)</td>
<td>'case'</td>
</tr>
<tr>
<td>gôn</td>
<td>ngang</td>
<td>(Fr. golf)</td>
<td>'golf'</td>
</tr>
<tr>
<td>(xanh) lơ</td>
<td>ngang</td>
<td>(Fr. bleu)</td>
<td>'blue'</td>
</tr>
<tr>
<td>phanh</td>
<td>ngang</td>
<td>(Fr. frein)</td>
<td>'brake'</td>
</tr>
<tr>
<td>xuya</td>
<td>ngang</td>
<td>(Fr. sûr)</td>
<td>'sure'</td>
</tr>
<tr>
<td>phi</td>
<td>ngang</td>
<td>(Fr. fût)</td>
<td>'barrel'</td>
</tr>
<tr>
<td>biu</td>
<td>ngang</td>
<td>(Eng.)</td>
<td>'bill'</td>
</tr>
<tr>
<td>meo</td>
<td>ngang</td>
<td>(Eng.)</td>
<td>'mail'</td>
</tr>
<tr>
<td>ba</td>
<td>ngang</td>
<td>(Eng.)</td>
<td>'bar'</td>
</tr>
<tr>
<td>phon</td>
<td>ngang</td>
<td>(Eng.)</td>
<td>'phone'</td>
</tr>
</tbody>
</table>

The predominant occurrence of the tone *ngang* in borrowings can also be seen in borrowings used as slang, e.g., hai-phai from English ‘hi-fi’ for ‘gay’, em-xi from English ‘emcee’, or in abbreviations, e.g., xi-di from English ‘CD’, ti-vi from English ‘TV’. Due to the dominance of *ngang* in borrowed forms in sonorant-final syllables, Avery 1983 states that *ngang* is the only tone in borrowings. However, although it occurs less frequently than *ngang*, *huyen* is also found. Interestingly, in (37), we find that borrowings that are monosyllabic words occur only with the tone *ngang*. In languages with stress, assuming that lexical monosyllabic words (e.g., verb, noun, adjective) bear stress, it would also be interesting to examine borrowings that are non-monosyllabic to see whether stress in the
original language is reflected in tones, i.e., whether the contrast is kept, but realized as tones.

The fact that only ngang or huyen occur in borrowed forms is used in many studies as evidence that those tones are unmarked. For example, Avery 1983 argues that ngang and huyen have no tonal tier; Burton 1992 argues that ngang and huyen are unmarked contour tones.

Hoi occurs in a very few items, e.g., ‘oăn (hoi) tù (huyen) ti (huyen)’ from English ‘one two three,’ or ‘mó (hoi) lét (sac2)’ from French ‘molette’. To my knowledge, borrowings with the tone nga are not found. The limited distribution of hoi and nga is not surprising: they are the most marked tones in the group that does not end with a stop consonant.

2.3.8. Summary

In this section I have argued for a structural representation of tones that incorporates markedness relations. Several types of neutralization that occur in reduplication, dialects, and poetry are predicted by the model, confirming markedness relations between the tones. Tones can be neutralized in the same register (e.g., nga > nang1 in Nghe Tinh dialects) or across register (e.g., nga > hoi in Southern dialects). Frequency of occurrence of tones in clitics also provides evidence for markedness. Finally borrowings show compatibility with the claim that ngang and huyen are least marked. The classification of tones presented above is based on the phonological patterning of tones (e.g., reduplication patterns and dialect neutralization) in a systematic way. In addition to phonological evidence for the general complexity of tones presented here,
phonetic evidence for the treatment of register is presented in Chapters Three and Five. The traditional literature bases the classification of tones on pitch as a register feature, while I have assumed here that register features are [laryngeal]. In the following sections, I discuss how tones are classified and represented in the traditional literature.

2.4. Traditional classification and representation of tones: the mismatch problem

In this section I provide an overview of the classification of Vietnamese tones and tonal features in the traditional literature. I will focus only on the classification and representation of tones in the traditional literature rather than on the inventory, which is discussed in Chapter Four. I argue that the traditional classification leads to a problem, namely it forces a mismatch between phonological patterning and the phonetic realization of tones in reduplication. I also show how this problem is accounted for in the literature.

2.4.1. Representations of Vietnamese tones in the traditional literature.

In this section I briefly describe different representations of Vietnamese tones in the literature. The phonetic features in these analyses will be discussed in detail in Chapter Five. In this section, I discuss only the phonological aspects of these analyses.

The usual representation of tones in the traditional literature is primarily a classification based on tonal shape and tonal height, e.g., which group of tones belongs to the low register, which group belongs to the high register, and which pairs of tones share a contour shape.
The assumption that register is equivalent to pitch is common in the Vietnamese literature. The term 'contour' in the traditional literature clearly refers to tonal shape. Tones are thus classified according to their pitch and shape. The classification of Vietnamese tones in Doan 1977, Vuong and Hoang 1994, and Hoang Thi Chau 1989, among others, is summarized in (38). The first parameter is the tonal shape or contour. According to this parameter, the term 'level' indicates that the tone is basically flat. The level tones are ngang and huyen. The non-level tones are sac1, nang1, hoi, and nga. Among these tones, nga and hoi are broken (the tone goes down then rises to make a curved shape); sac1 and nang1 are non-broken (the tone either goes down or rises) (see Chapter Three for details). The view of contour is consistent with my own. The second parameter is tonal height, which divides tones into two registers, defined by pitch: high register consists of the high tones: ngang, sac1 and nga, and low register consists of the low tones: huyen, nang1 and hoi. The justification of [high] and [low] in (38) is discussed in Chapter Three where I review the acoustic experiments in the literature.

(38) Traditional classification of tones according to tonal shape and height

<table>
<thead>
<tr>
<th></th>
<th>level</th>
<th>non-level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>non-broken</td>
</tr>
<tr>
<td>High</td>
<td>ngang</td>
<td>sac1</td>
</tr>
<tr>
<td>Low</td>
<td>huyen</td>
<td>nang1</td>
</tr>
</tbody>
</table>

Evidence for the two registers high and low is drawn from phonetics (e.g., Doan 1977, Vuong and Hoang 1994 among others). Evidence for tonal contour comes from tonal patterning in poetry, i.e., level tones form one group and non-level tones form the other
group (See section 2.3.4). These sources are briefly mentioned as evidence for classification, but there are no careful or systematic accounts given. However, when we look at how tones are classified according to their pattern in reduplication, the situation is different.

The classification in (38) is standard in the traditional literature. The terms in (38) are used primarily to classify tones, not to specify features of tones.

Let me now turn to the most significant problem for the classification in (38). This differs in one major way from my proposal, repeated in (39).

(39) Proposed classification of tones

<table>
<thead>
<tr>
<th>Unmarked register:</th>
<th>ngang</th>
<th>sac1</th>
<th>hoi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marked register:</td>
<td>huyen</td>
<td>nang1</td>
<td>nga</td>
</tr>
</tbody>
</table>

The major difference between the traditional classification in (38) and my proposal in (39) comes in the tones identified as broken: in my classification, based on phonological patterning, hoi forms a class with ngang and sac1 in terms of register and nga forms a class with huyen and nang1. In the traditional proposal, the registers of hoi and nga are reversed.

Tonal features are also elaborated. In the literature, besides pitch height, other features such as glottalization, length, tense and lax are proposed as features of tones, although evidence is not given. For example, (40) shows the tone features from Thompson 1965, with the feature for tonal length from Alves 1997. Again, (40) assumes a six-tone-view.

56
(40) Tone features from Thompson 1965 and Alves 1997

<table>
<thead>
<tr>
<th>Tone</th>
<th>high</th>
<th>tense</th>
<th>glottalic</th>
<th>short</th>
</tr>
</thead>
<tbody>
<tr>
<td>nang</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>huyen</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>sac</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>nang</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>hoi</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>nga</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

In (40), tonal height is also used to classify tonal registers. One problem with (40) is that these features do not group tones into natural classes. For instance, ngang, sac and nga are grouped under the feature [high], but they do not pattern together as a group. In addition, although nang and nga have [glottalic], they do not pattern as a pair.

Hoang Cao Cuong 1986 argues that two parameters alone, pitch height and contour, cannot account for tone languages such as Vietnamese, whose tones developed from final consonants (see Chapter Four for details) and still show some relic of those segments through glottal stop or breathiness. In an attempt to incorporate phonation types as tone features, Hoang Cao Cuong 1986 proposes the tonal features in (41). This analysis classifies tones based on a diachronic hypothesis, i.e., that tones developed from final consonants (see Chapter Four). In this analysis, a tone that developed from a final consonant may or may not retain some properties of that segment. For instance, the glottal stop in the tone nang1 is regarded as a relic of the historical final glottal stop. In (41), the phonation column indicates the presence or absence of glottalization or
pharyngealization in a tone. The second column indicates whether the tone falls within its register, i.e., pitch contour does not cross two registers. The last column indicates the pitch height for register features. Note that Hoang classifies eight tones according to their phonetic properties, but the phonological status of the sac and nang tones is not clear in this analysis.

(41) Tone features from Hoang Cao Cuong 1986

<table>
<thead>
<tr>
<th>Tone</th>
<th>phonation</th>
<th>falls in one register</th>
<th>low</th>
</tr>
</thead>
<tbody>
<tr>
<td>ngang 1</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>huyen 2</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>nga 3</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>hoi 4</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>sac 5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>nang 6</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>sac 5'</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>nang 6'</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

The features in (41) are phonetically classified. Although Hoang argues that the classification of tones in (41) agrees with the traditional treatment in music and in frequency of occurrence of tones in the reduplicative forms, evidence is not provided.\(^{11}\)

---

\(^{11}\) Hoang states that in music, if a tone does not change much in F0, but still has some phonation and falls within its register, it is not regarded as an even tone (vs. non-even tone). No further evidence is provided. In reduplication, he states that reduplicative forms with tones that share phonation and even/ non-even features occur more than reduplicative forms with other tones. The argument does not go further than this statement.
However, he emphasizes that (41) is primarily based on phonation, which shows historical developments of tones. Most importantly, like the classification in (40), these features do not account for the major patterning of tones in the language, as we have seen in section 2.3, e.g., nga, nang1, sac2 and nang2 do not pattern as a group although they share the same feature for phonation according to the classification in (41). Hoang recognizes that it is problematic to use pitch height to classify Vietnamese tones, because while some tones fall completely in one register (high or low), other tones fall across the two registers. This issue is further discussed in Chapter Five.

All the featural accounts discussed in this section fail in one major way: the proposals do not capture classes that the phonology requires to be natural. Other recent works use feature geometry to represent tones as constituents, but do not provide evidence for the organization of tones. For instance, based on the features in (41), Hoang Cao Cuong 1986 suggests the feature geometry in (42) for Vietnamese tones. The numbers here refer to tones as follows: 1 is ngang, 2 is huyen, 3 is nga, 4 is hoi, 5 is sac1, 5' is sac2, 6 is nang1, and 6' is nang2. The abbreviation 'ng' stands for ngang, 'hu' for huyen, 'sa1, sa2' for sac1 and sac2, respectively, and 'na1, na2' for nang1 and nang2, respectively.
Feature geometry for Vietnamese tones in Hoang 1986

However, these features do not group tones in the way that they pattern in the language. For instance, there is no evidence that 3 6 5' 6' (nga, nang1, sac2, nang2) on the top row pattern as a group; or 5 and 4 (sac1 and hoi) on the last row pattern together as a pair, and so on.

Note that in classifying tones, it is problematic to use only phonetic features such as length, glottalic, tenseness and so on. For instance, the feature [glottalic] groups nga and nang1 into a class which does not reflect their phonological patterning. In an attempt to avoid the difficulty with the feature [glottalic], Ngo Thanh Nhan 1984 suggests another feature, one which makes use of the tonal shape, namely [concave], to group hoi and nga together. In this system, tonal representation is simply a reflection of the pitch shape of tones, or their contour (1984:75). Features of tones in this account are shown in (43).
(43) Tonal features Ngo (1984)

<table>
<thead>
<tr>
<th></th>
<th>ngang</th>
<th>huyen</th>
<th>sac</th>
<th>nang</th>
<th>hoi</th>
<th>nga</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concave</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Contour</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>High</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Using the tonal features in (43), Ngo Thanh Nhan represents tones hierarchically as in
(44) with a binary branching model. This model uses [concave] as the highest level
which dominates tonal contour. The feature [±high] is a dependent of [contour]. The
high tone hoi, for instance, is [+concave, +contour, +high]. The low tone nga is
[+concave, +contour, -high]. According to Ngo, this is a morphological representation,
since hoi is phonetically low but morphologically high, and nga is phonetically high but
morphologically low.

(44) Feature geometry for Vietnamese tones in Ngo 1984

```
+concave
- -
+contour
- -
\hline
-hi
+hi
- -
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hliner
There is an unexplained gap in the model: the absence of two tones under [+concave, -contour]. This gap is also seen in Doan (1977:107), who explains that in reality there are no tones which are ‘even’ but ‘broken’ (i.e., [- CONTOUR] and [+CONCAVE] in Ngo 1984, respectively). Using binary branching with three different levels, the model seems to be set for an eight-tone system (i.e., combinations give eight possible terminal nodes), but it is not completely filled. We would expect the two tones ngang and huyen to occupy the two unspecified slots as their unmarkedness is agreed upon among authors (see section 2.3, Burton 1992, Avery 1983, Ngo 1984), based on their frequency, their patterns in reduplication and borrowings. However, these two tones do not occur there. Most important, this model does not capture the distributional fact that the four tones ngang, huyen, hoi and nga only occur in sonorant-final syllables.

In summary, although these analyses classify tones into pairs, there are no relationships between those pairs, therefore, tonal pairs exist as autonomous components in the tonal system. Accordingly, there are no adequate representations that can account for the various patterns of tones in the Vietnamese language. The division of tonal registers in (38), and the assumption that register is tonal height, also cause a serious mismatch between the phonetics and phonology of tones. The next section will examine this problem in detail.

2.4.2. Mismatch between the phonetics and phonology in reduplication.

The phonetics-based classification of tones in (38) creates a problem in reduplication with respect to the two tones hoi and nga.
As we have seen in section 2.3.1, tones pattern in two registers: \textit{ngang} and \textit{sacl} pattern together, and \textit{huyen} and \textit{nang} pattern together. \textit{Ngang} and \textit{sacl} are high tones while \textit{huyen} and \textit{nang} are low tones. The phonetics of these tones predicts their patterning, i.e., their phonetic classification underlies their phonological classification. Given this patterning, one would predict that \textit{nga}, a phonetically a high tone, should pattern phonologically with \textit{ngang} and \textit{sacl}, and \textit{hoi}, a phonetically low tone, should pattern with \textit{huyen} and \textit{nang}. However, when we look at \textit{hoi} and \textit{nga}, they do not confirm the pattern predicted by the phonetics. In (39), repeated here as (45), \textit{hoi} patterns phonologically with \textit{ngang} and \textit{sacl}, as if it were a high tone. \textit{Nga} patterns phonologically with \textit{huyen} and \textit{nang}, as if it were a low tone.

(45) Phonological patterns of tones in reduplication

<table>
<thead>
<tr>
<th>Unmarked</th>
<th>\textit{ngang}</th>
<th>\textit{sacl}</th>
<th>\textit{hoi}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marked</td>
<td>\textit{huyen}</td>
<td>\textit{nang}</td>
<td>\textit{nga}</td>
</tr>
</tbody>
</table>

Many efforts have been made to explain the unusual behavior of \textit{hoi} and \textit{nga} in reduplication. Some researchers (Doan (1977:123), Hoang Chau (1989:207), Vuong & Hoang (1994:100), Nguyen and Edmondson 1997 among others) look at historical developments as an explanation for this phenomenon (see Chapter Four). They adopt Haudricourt's suggestion that \textit{hoi} was historically a low tone and \textit{nga} a high tone, and they suggest that \textit{hoi} and \textit{nga} switched their registers during the evolution of tones. How this happened is not clear, nor is it closely examined in most accounts. Ngo 1984 and
Burton 1992 propose abstract representations of tones to solve the problem with *hoi* and *nga* in reduplication. I first present Ngo’s analysis.

In order to account for the phonetics-phonology mismatch with *hoi* and *nga*, Ngo posits a Concave Tone Reversal rule which changes the phonological high tone *hoi* to a phonetic low tone and the phonological low tone *nga* to a phonetic high tone (1984:78). The rule he proposes is shown in (46).

(46) Concave Tone Reversal

\[
\begin{align*}
\text{t} \\
[\alpha \text{ high}] & \rightarrow [-\alpha \text{ high}] \\
[+\text{concave}]
\end{align*}
\]

This flip-flop rule does not have any theoretical motivation. This type of rule is not found in other systems synchronically: for instance, we do not find languages where a phonetically back vowel patterns phonologically as if it were a front vowel and vice versa. Such a system would necessitate a flip-flop rule of the type in (46). The absence of this rule type makes (46) suspicious, and to be avoided if possible.

I now turn to Burton’s solution. In terms of the behaviour of *hoi* and *nga* in reduplication, let us first look at how Burton classifies these two tones. Because of the discrepancy between (38) and (39) involving *hoi* and *nga*, other classifications have been proposed. Based on the pattern of tones in reduplication, reflected in the classification in (39), Burton 1982 proposes the representation in (47) for Vietnamese tones. Adopting Yip's 1980 model, Burton uses upper case +HIGH/ -HIGH for Register and lower case

64
h, l, lh for contour features. In this model, each tonal pair shares the same feature for contour, e.g., both ngang and huyen are 'l'. They differ only in the register feature, i.e., +/- HIGH.

(47) Representation of Vietnamese tones (Burton 1982)

<table>
<thead>
<tr>
<th>REGISTER</th>
<th>CONTOUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ HIGH</td>
<td>A = l (ngang)</td>
</tr>
<tr>
<td>- HIGH</td>
<td>D= l (huyen)</td>
</tr>
</tbody>
</table>

Note that Burton adopts the view that sac2 and nang2 are variations of sac1 and nang1, respectively, in stop-final syllables, see Chapter Four. Compared to the table in (38), Burton’s classification is abstract in that it groups hoi, a phonetic low tone, with phonetic high tones, and nga, a phonetic high tone with phonetic low tones. However, at the same time, he also claims that ‘Upper vs Lower pitch-range is exactly what characterizes Register’. Similarly, Burton classifies ‘contour features’ using both phonological and phonetic properties. This issue is discussed in detail in Chapter Five.

With respect to tonal contour, hoi and nga are rising ‘lh’ based on the theory of tones that he adopts (Yip 1980, Bao 1999b, Duanmu 1990). This theory does not allow a dipping tone, i.e., ‘hh’. Burton treats the initial dip in these tones as a phonetic effect. However, with respect to register, while arguing that ngang and sac are [+High] register and huyen and nang are [-High] register because they are articulated in the upper and lower part of their pitch-range, respectively, Burton does not say why hoi is [+High] and

65
 nga is [-High]. We can guess that he does this to account for the patterning of tones in reduplication.

Now let us consider how Burton accounts for the mismatch between hoi and nga. He proposes that hoi and nga must switch their register to achieve the appropriate phonetic representations. He argues for the surface realization of hoi and nga in Northern Vietnamese as follows. The surface representations of hoi and nga result from an interaction of two rules: Creakiness Acquisition and Register Flip-Flop in (48).12

(48) a. Flip-Flop rule
   lh[αHigh]       -->          lh [-αHigh]

b. Creakiness Acquisition:
   Low-register h   -->      [+creaky]

Recall that in order to capture the fact that hoi patterns phonologically with high register tones and nga with low register tones, other authors suggest that diachronically hoi was a low tone, which changed to a high tone. They must then assume that the phonology of reduplication does not make reference to natural classes. Unlike those authors, Burton argues that hoi is synchronically [+High]. He explains the phonetic surface of hoi as a high tone by a Flip-Flop rule which switches its register. Similarly, nga is underlyingly [-High] but it surfaces as [+High] by a Creakiness Acquisition rule (to get the phonetic feature [creaky]) followed by a Flip-Flop rule (to switch its register).
Register Flip-Flop changes \textit{hoi} from an underlying \([+\text{High}]\) to surface \([-\text{High}]\). Thus, Burton’s Flip-Flop rule is equivalent to Ngo’s Concave Tone Reversal in (46); the major difference between these analyses is simply whether the tonal switch is considered to be diachronic or synchronic.

In summary, we have seen that features of tones have been largely phonetically based in the traditional Vietnamese literature. These features do not group tones into natural classes. However, when one looks at patterning in reduplication and neutralization, a classification of tones which relies only on a phonetic interpretation of register as high and low is not a satisfying one, since it leads to the mismatch problem between the phonetics and phonology of the tones \textit{hoi} and \textit{nga}. The analyses based on this type of classification account for the mismatch with an unnatural rule that switches the registers of these two tones.

2.5. Summary

In this chapter I have proposed a structural representation of Vietnamese tones and evidence for it. I have also reviewed the representations of tones in the literature, and the mismatch problem that they create. This problem arises to a large degree, I believe, because of the treatment of register as involving pitch height.

\footnote{Burton proposes the Creakiness rule to insert this feature phonetically rather than have it present phonologically, arguing that if the feature \textit{creaky} (in the tones \textit{nga} and \textit{na\text{ng}i}) were treated as a distinctive feature, the number of possible tones would be doubled in a given language.}
Given the unnaturalness of a flip-flop rule as a synchronic process, one must ask if an alternative analysis is available. Given the phonetic classification in (38) and the phonological system in (45), it is difficult to imagine any alternative analysis.

While register is assumed to involve the height of F0, one also finds statements in the literature about tones involving breathiness and creakiness. Given that a pitch-based analysis of register is problematic, I decided to carry out a detailed investigation of the tones in Northern dialects to see whether pitch height indeed provides the most appropriate phonetic basis for classification into registers. To put this another way, I examine whether what I have labelled as unmarked and marked (see (20)) should be labelled high and low pitch, as in the literature, or in terms of laryngeal features, as proposed (without argument) in (18). In the following chapter, I present my experimental study and its results.
Chapter 3

AN ACOUSTIC STUDY OF THE TONES OF THE HANOI DIALECT

In this chapter, I first review previous experimental studies of Vietnamese tones and then present an acoustic study with Northern dialects. In particular, I discuss fundamental frequency, phonation types, length and the linear portion of tones.

3.1. Literature: acoustic studies

In this section, I summarize only the results of previous acoustic studies on Vietnamese tones. Scattered observations about tones in the literature on Vietnamese phonology are not mentioned.

There have been a number of acoustic studies of Vietnamese tones (see Mieko Han 1969, Vu Thanh Phuong 1982, Hoang 1986, Nguyen and Edmondson 1997, Vu Ba Hung 1999, among others). Here I review Vu 1982, Hoang 1986, and Nguyen and Edmondson 1997, three studies with graphs showing Fo of pitch, together with the phonetic descriptions of tones as given in every study in the traditional literature.

3.1.1. Nguyen and Edmondson 1997

Nguyen and Edmondson (NE) 1997 is the only study using modern instrumental techniques that focuses on phonation types in Northern Vietnamese tones. NE argue that in addition to F0, voice qualities are used distinctively in Northern dialects.
NE examined an open syllable, [ta], with 6 tones; tones in stop-final syllables were omitted. Six speakers were recorded three times each and the responses were analyzed using CECIL 2.1. The F0 value of each syllable was calculated. Figure 1 in Chapter 2 is repeated below to show the normalized F0 values of the three pairs of tones. It is not clear whether the values given in Figure 1 are averaged across speakers or from one speaker. Note that the semitone scale is used to measure F0 in this research instead of frequency in Hertz.

Figure 1. F0 values of three tonal pairs in Nguyen and Edmondson (1997:6).

Fig. 1a. Ngang and huyen in [ta]

Fig. 1b. Sac1 and nang1 in [ta]
The following description is provided by the authors. Each tone is given a relative pitch value in numbers using the system in Chao 1930 (cited in Bao 1999b), with 1 representing the lowest point and 5 the highest point in the pitch range. The first pair in (1a) is ngang and huyen. Ngang (high level) is a high tone with a flat contour gradually going down a little bit. It is given the value of 33. Huyen (low level) is a low tone with a flat contour, starting lower than ngang and going slightly down at the endpoint. It is given the value of 21. The second pair in (1b) is sac1 and nang1. Sac1 (rising) starts from a lower point than ngang, going up to the highest point in the pitch range. It is a 35 tone. Nang1 (falling) is shorter than the other tones. It is a 32 tone. The third pair in (1c) is nga and hoi. Nga starts as high as ngang and then goes up to the highest point in the pitch range. It is interrupted by a glottal stop at approximately 225 ms. Hoi starts between the beginning points of ngang and huyen, goes down, and then rises to a point that is close to the onset. NE do not give tonal number values for hoi and nga.

Nguyen and Edmondson 1997 observe that there is variation in tones from speaker to speaker in terms of length, contour and F0 level. Figure 2 gives a general
picture of the six tones using the syllable [ta] from two speakers. Note that in Fig. 2a, the
sac tone in speaker 6 starts higher than ngang, huyen, and hoi, at approximately 44
semitones. However, in speaker 2 in Fig. 2b, sac has the lowest starting point. The nang1
tone is very short in speaker 6. More interestingly, it is higher than the same speaker's
ngang. This point is important for my analysis, and I will return to it later, in Chapter
Five. Also note that, as in Figure 1, Figure 2 shows that nga for both speakers is broken in
the middle by a glottal stop, and the glottal stop for both speakers is longer than that in
Figure 1, as NE point out. In particular, in speaker 6 in Fig. 2a, nga starts at 46 semitones
and is broken by a glottal stop from 100 ms to approximately 320 ms. After the glottal
stop, the rising portion emerges between 48 and 50 semitones. For speaker 2 in Fig. 2b,
the broken portion in the nga tone is shorter than that of speaker 6, from approximately
90 to 180 ms. After the glottal stop, the rising portion in Fig. 2b emerges between 50 and
54 semitones, higher than that of speaker 6 in Fig. 2a. The hoi tone does not show any rise
in either speaker.
Figure 2. Six tones from speakers 2 and 6 in Nguyen and Edmondson (1997:8)

Fig. 2a. Six tones from speaker 6 (NE 1997)

Fig. 2b. Six tones from speaker 2 (NE 1997)
NE focused on the phonation types found in the Northern dialects. Their results will be discussed in Chapter Five, section 5.1.2.

3.1.2. Vu Thanh Phuong 1982

This study investigates the general F0 of tones. Vu 1982 recorded eleven speakers from Northern dialects reading word lists. One speaker was from Thanh hoa province and one was from between Nam dinh and Thanh hoa.\(^\text{13}\) Vu also included the two tones that occur in stop-final syllables, \textit{ sac}2 and \textit{nang}2. The diagrams in Figure 3, from Vu, show the mean F0 of eight tones plotted against mean duration (in centiseconds) from eleven speakers.

\(^{13}\) According to Hoang Thi Chau (1989:206), Thanh Hoa speakers do not distinguish \textit{hoi} and \textit{nga}.
In Figure 3, the two tones that do not appear in NE are sac2 and nang2. Sac2 is the highest tone, short and rising steadily to the highest point. Nang2 is the second lowest tone, short and falling steadily. If we ignore sac2 and nang2, we can see the general similarities between the F0 patterns of the tones in Vu and those in NE. For instance, sac1 and nga have the highest end points; nga is broken by a glottal stop (represented by a dotted portion). However, Vu shows nga as going down steadily and then rising very quickly to form a V shape. The apparent difference in F0 between tones in Vu is larger than in NE due to the different scales used. Unlike NE, Vu’s nang1 is much lower than
ngang. Notice also that the curved tone hoi, the lowest one in Vu, clearly has a rising part, which is not present in NE. Therefore, although the two studies agree on many points, they show variation on other points such as the length of the glottal stop or of nang1, or the shape of the curved tone hoi.

3.1.3. Hoang 1986

The results of Hoang 1986 show tonal shapes in the Hanoi dialect that are similar to those seen in Vu 1982. However, we also find the differences in F0 between these two studies. Hoang recorded eleven speakers reading a word list of 150 syllables. Figure 4 shows pitch graphs of eight tones from four speakers. Speakers 1 and 2 are male; speakers 3 and 4 are female. All tones are labeled in the diagram for Speaker 4. Hoang does not show duration.
As in Vu 1982, for Hoang the three tones with the highest ending points are sac1, sac2 and nga. Nga also has the clear V shape found in Hoang, but the break for the glottal stop is not shown. Let us look closely at the two tones hoi and nga which show variation between speakers. Hoi is represented by a line broken by dots. It is the lowest tone in Speaker 1; it has a rising part in Speaker 2; it falls and levels off in Speakers 3 and 4. Nga
is represented by a broken line with small circles. Nga falls and rises. The most noticeable difference between the two curved tones hoi and nga in this study and in the studies of NE and Vu is that, for NE and Vu, the lowest point of nga is higher than that of hoi, while for Hoang hoi bisects nga, as with speakers 3 and 4. This is also a crucial point to be discussed in Chapter Five.

3.1.4. Summary

From these studies, we can see variation in both the pitch height and contour of tones, especially the non-level tones. However, it is important to note that all authors describe nga as having a glottal stop in the middle, and all share the same view on how tones are classified according to their contour and registers.

Table 1 shows the classification of tones that arises from these phonetic studies. Note that in these studies (and in almost all traditional studies of Vietnamese tones), ‘register’ is to be interpreted as ‘tonal height’. High register tones include ngang, sac1, nga, and sac2. Low register tones include huyen, nang1, hoi, and nang2.

<table>
<thead>
<tr>
<th></th>
<th>ngang</th>
<th>sac1</th>
<th>nga</th>
<th>sac2</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>huyen</td>
<td>nang1</td>
<td>hoi</td>
<td>nang2</td>
</tr>
</tbody>
</table>

This division causes a serious problem involving the phonetics and phonology of tones, which is discussed in Chapters Two and Five.
3.2. Experimental design

This study investigates F0 and the phonation types found in the tonal system of Northern dialects. It also examines how these features are measured and evaluated.

My acoustic study investigated the fundamental frequency (F0) of Vietnamese tones and the spectral characteristics of phonation types such as creakiness and breathiness. Some aspects, including amplitude and individual differences, are left for future research. Specifically, I examine the F0 of the eight tones, and any phonation differences they might have in the acoustic signal.

3.2.1. Data and subjects

In this study, speakers of Vietnamese were asked to read controlled lists of words. Each word consists of a CV or CVC syllable. The syllables are described below. Vowel length and voicing of initial consonants were not included in the data (see Appendix 4). Only one vowel, [a], was used. The vowel [a] was chosen because other vowels are usually diphthongized in different dialects. The initial consonants in open syllables include a nasal /m/, two stops /t, k/, and two fricatives /s, z/. The final consonants include stops and velar nasals. The two final glides were tested only with the two tones hoi and nga. Some other forms were mixed in with tested forms. These included the long vowel /a:/ with the final nasal /m/, (tắm tăm, tậm tần), the short vowel /a/ with the final nasal /n/ (tần tần, tần tần), and closed syllables with the diphthong /iə/. The set of tested syllables thus includes the segmental shapes [ma], [ta], [sa], [za], [ka], final consonant /ŋ/ and stops plus appropriate tones.
Recall that there is no tonal sandhi in Vietnamese. All tested forms were put in a fixed elicitation frame. Therefore, the effect of the elicitation frame should be consistent for all data. The frame sentence is Chữ...phải ngay [củ ......fa:j ɲaj] 'The word...has to be straight'. The fricative /f/ was chosen to follow the target form because there is a clear boundary between the target form and /f/ in the spectrogram, which is indicated by noise at higher frequencies.

The study involved nine speakers, three males and six females, of the Northern dialects (see informants' information in Appendix 3). They are between 17 and 49 years of age and live in Toronto. They come from Ha Noi, Hai Phong and Nam Dinh (see map in Appendix 1). Six speakers were born and grew up in the North (Hung, Hoang, Phuong, An, Binh and Dung). Three speakers were born in the North and spent their childhood there. Later they left for the South, but they lived with their family of the same background and within Northern communities (Khanh, Son, Van). These people speak Northern dialects and regard the Northern dialect as their native dialect. Depending on individuals, they have been in Toronto from one month to about eighteen years. All speakers were completely unaware of what I was testing for in their speech.

There are 112 items randomized 3 times and presented to each speaker individually. The speaker read each of the 3 lists. With nine speakers, the total number of tokens is 3,024 (See Appendix 4).

3.2.2. Procedures and methodology

The spoken responses of subjects were recorded in a quiet room. They were then digitized on an iMac computer. Each data file contains the target form with its
surrounding environment, i.e., after an open syllable [cu] and before [f], the onset of the following syllable, [fa:j]. Signalyze\textsuperscript{TM}3.12, created by Eric Keller 1994, was used for the analysis. Fundamental frequency was not measured by pitch-extraction from the program but was manually measured by determining the length of the cycle at a 30 ms interval. The signal was digitized. The raw signal then was enlarged in order to see clearly individual cycles. Signalyze was also used to produce spectrograms. Excel was used to graph pitch. In addition to F0, both creakiness and breathiness were clearly present in the waveforms and they were examined as well (see characteristics of breathiness and creakiness in section 3.4).

F0 was measured every 30 ms, starting immediately after the burst of an initial stop or after an initial nasal segment. Measurements were taken at 30, 60, 90, 120, 150, 180, 210, 240, and 270 ms. For the tones with creakiness or glottal stop, i.e., nang\textsuperscript{i} and nga, F0 was measurable only when there was a periodic glottal pulse evident in the spectrogram.

With breathy tones, F0 was measured when individual cycles could be clearly identified. When the tone is heavily breathy (i.e., glottal pulses with reduced amplitude, it is very difficult to identify individual cycles), the recorded signal was filtered using a low pass (0-500Hz) Butterworth filter to suppress the high frequency components. After filtering, individual cycles were more easily seen, with only low frequency components preserved. If there was no voicing in the breathy portion, F0 was unmeasurable, but this was rare.

With a nasal-final syllable, the second part of the tone is realized during the nasal. Although voicing is not strong in the nasal portion of the spectrogram, F0 was still
measurable. If this part is ignored, the tone loses its second part and becomes unusually short; In such cases, there is not enough information about the tone's contour, and the tone becomes unrecognizable. Usually the first two tokens of a form were measured. The third token was examined only as required, e.g., when the speaker made an unusual pronunciation of a form due to a slip of the tongue or had an unusually long pause before or after the form.

A spectrogram of the signal was produced using Signalize's extra-wide (300Hz) setting for a female voice and very wide (200Hz) for a male voice.

3.3. Results

In this section I present my findings on various aspects of tones: the fundamental frequency, phonation types, length and linear portion of tone.

3.3.1. Descriptions of fundamental frequency of eight tones

I will first show the values of F0 across time, measured from the raw waveform (not the filtered waveform). The characteristics of tonal shapes provide evidence for the contour feature to be discussed in Chapter Five.

3.3.1.1. Fundamental frequency

In this section I provide a description of the F0 contours of six tones with the same syllable [ta], and of sac2 and nang2 with [ta:k], from eight speakers. Graphs showing the F0 contours of eight tones from three males and five females are given in Figure Five.
The values of tokens of the same form in each speaker as well as across speakers are not averaged.

The legend of tones in the graph is: the speaker's name, followed by the repetition of the token, then the tone: 'ng' is ngang (high-level), 'hu' is huyn (low-level), 'sac' is sac1 (high-rising), 'na' is 'nang1' (low-falling), 'sac-cac' is sac2 (high-rising in stop-final syllables [ka:k]), 'na-cac' is nang2 (low-falling in stop-final syllables), 'hoi' is hoi (falling-rising), 'nga' is nga (falling-rising with a glottal stop). For instance, in Figure 5a, VKhi.ng-ta means speaker VanKhanh, first token, tone ngang, syllable [ta]. Note that the form in the legend box represents the writing system, e.g., ta is the written form of [ta], da is the written form of [za]. The phonetic form is used in the title of the graph, e.g., '6 tones-za' means the pitch graph (F0 graph) of six tones with the form [za].

In Figure 5, the three male speakers Hung, Hoang and Son always have the lowest frequencies. Of the three males, Hung has widest frequency range. Hoang is in the middle of the range except in sac1 (5c) and sac2 (5g). I now provide descriptions of the general characteristics of each tone.
Figure 5a. Tone *ngang* from eight speakers

*Ngang* (5a) is level and usually falls slightly during the last 30 ms. *Ngang* starts at about 210 to 270Hz in females and 120 to 175Hz in males.
Figure 5b. Tone *huyen* from eight speakers

*Huyen* (5b) is also level and slightly falling. *Huyen* starts around 150 to 250Hz in females and 110 to 160Hz in males.
Figure 5c. Tone sac1 from eight speakers

Sac1 (5c) starts slightly lower than huyen, at about 180 to 230Hz in females and 100 to 150Hz in males. It is mainly level, rising during the last 60ms. The length of the level part varies from speaker to speaker. In some speakers (An and Hoang), the level part is quite long at approximately 150ms. However, generally after about 120ms, this tone rises and ends between 220 to 320Hz in females, and 120 to 190Hz in males. Depending on the speaker, the F0 difference between the beginning and end points of this tone is about 40 to 90Hz in females, and 20 to 40Hz in males. The height of the ending of sac1 varies from speaker to speaker, e.g., around 250Hz in VKhl but up to 320Hz in BinhI. It also varies from form to form in the same speaker.
Figure 5d. Tone *nang*₁ from eight speakers

![Graph showing frequency over time for eight speakers.](image)

*Nang*₁ (5d) starts between about 180 to 280Hz in females, and 100 to 155Hz in males. It falls steadily after approximately 90 ms for most speakers. It has either a glottal stop (a blank in the line showing pitch) or a very creaky portion from 150ms to 180ms where it looks very chaotic in the pitch graph. Then the tone remains level or rises slightly before the end. Whether the tone ends with a glottal stop or keeps going for a while varies from speaker to speaker and from token to token.
Figure 5e. Tone hoi from eight speakers

Hoi (5e) starts the same as the onset of ngang or nang1 and gradually falls to about 140 to 180Hz in females and 90 to 120Hz in males. Then it rises slightly again. In some cases, the rising part is absent, and the tone continues to fall lower than huyen.
Nga (5f) usually starts higher than hoi, around the onset of huyen, and falls steadily during the first 90ms. At this point, it is either broken by a glottal stop or it becomes very creaky for 90 to 120ms. If the tone is broken by a glottal stop, the pitch curve is interrupted in the middle of the tone. If it is creaky, the lowest point in nga is around 70Hz (in PhuongI). After the creaky part, the tone rises. The height of the ending varies from speaker to speaker and from token to token.
Note that in order to examine sac2 and nang2, it is necessary to use obstruent-final syllables.

Sac2 (5g) has the highest starting point and is short. It rises and ends after around 150ms in females and 100ms in males. This difference between female and male speech could be a sociolinguistic marker because a long sac2 sounds more feminine.
Figure 5h. Tone nang2 from eight speakers

Nang2 (5h) is also very short. This tone is very similar to huyen from beginning to end, but much shorter.

Figure 6 shows six tones for [ka] along with sac2 and nang2 for [ka:k] from a female speaker (see Appendix 5 for other speakers). The title in each graph from one speaker is to be read as follows: the speaker’s name is followed by the repetition of the token, then the tone and the form. For example, in Figure 6, the title ‘Binh I. 8 tones- ta’ is read as follows: the speaker Binh produced 8 tones, the first token of the form [ta]; ‘Binh II. hoinga-ka’: the speaker Binh produced the second token of hoi and nga with the form [ka].
In Figure 6, from Speaker Binh, we see that the range of F0 is between 50 and 350Hz, with the highest point in sac2, i.e., BinhLsac-cac, and the lowest point in nang1, i.e., BinhL.na-ca. We see here a variety of tonal shapes: tones are flat (level), go up (rising), go down (falling) or go down and then up (curved). However, if we look at the three lowest curved tones, i.e., hoi, nang1 and nga, it is difficult to say whether they all are always curved and only different in F0.

Figure 6. Pitch graph of eight tones in [ka]/[kaːk] from Speaker Binh

![Pitch graph of eight tones in [ka]/[kaːk] from Speaker Binh](image)

From the results above, tones can be classified by their contour: rising tones, falling tones and falling-rising tones, i.e., curved. For the moment, I omit the parameter of tonal height; I will discuss it in detail in the following section. Table 2 summarizes the...
findings about contour. This table is similar to Table 1, except that ‘high’ and ‘low’ have been replaced by ‘unmarked’ and ‘marked’ (see Chapter Two, section 2.2.2), and hoi is classified as an unmarked tone, and thus in the same category as ngang, sac1 and sac2. Nga is a marked tone belonging to the same category as huyen, nang1 and nang2 (see Chapter Two, section 2.2.2). In Table 2, tones are divided into two categories: even (level) tones, and non-even tones (tones that involve a change in fundamental frequency). The latter then is divided into three types of tonal shape.

Table 2: Classification of tone according to their contour

<table>
<thead>
<tr>
<th></th>
<th>EVEN</th>
<th>NON-EVEN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>rise/ fall</td>
</tr>
<tr>
<td>UNMARKED</td>
<td>ngang</td>
<td>sac1 (rising)</td>
</tr>
<tr>
<td>MARKED</td>
<td>huyen</td>
<td>nang1 (falling)</td>
</tr>
</tbody>
</table>

3.3.1.2. Remarks on tonal shape

Certain points need to be made concerning variations in the tonal shape of the two curved tones hoi and nga, and the two falling tones, nang1 and nang2.

3.3.1.2.1. HOI and NGA

Generally hoi goes down and then up slightly. However, each study differs in how curved this tone is. For instance, in Vu 1982 (Figure 3) hoi is reported to be very
curved, but in Hoang 1986 (Figure 4) the tone is said to go down and then slightly up.
Vu Ba Hung 1999 observes that the rising part is clearer and stronger in males than in females. My study also shows variation in the rising part of the tone hoi from speaker to speaker. For instance, in Figure 7, hoi is curved in Figure (7a), but not in Figure (7b).

Figure 7a. Hoi with a curve in [ta] from Speaker Binh

I include nga in this graph to show the glottal stop in the middle of the tone, and the low ending, which I will discuss next.
Figure 7b. *Hoi* without a curve in [ta] from Speaker An.

The curve in *hoi* does not appear consistently even between tokens from the same speaker using the same form. In Figure 7a of Speaker Binh, *hoi* is curved in the first token of [ta], BinhI-ta, but not in the second token, BinhII-ta, as seen in Figure 8. In Figure 8, *hoi* goes down and flattens out. Again, *nga* is included to illustrate the curved shape of the tone, and the high ending, which will be discussed next.
Figure 8. *Hoi* without a curve in another token of [ta] from Speaker Binh

**BinhII.hoinga-ta**

With *nga*, the tone is not necessarily broken by a glottal stop. It can have a creaky portion instead. Figures 7a and 8 show *nga* broken by glottal stop, while in Figure 7b, no glottal stop is present. This variation is also noted in Vu 1999. If the tone is broken by a glottal stop, the glottal stop is indicated by a discontinuous portion in the pitch curve. If the tone is creaky instead, the fundamental frequency is measured only when it is clear enough to see the individual cycles. In all studies including mine, *nga* always falls and then rises to make a clear curve. However, unlike the common description found in previous studies (Vu 1982 in Figure 3, Hoang 1986 in Figure 4), I did not find that this
Tone always has a very high ending for all speakers. This can also be seen in Figures 7a, 7b and 8. *Nga* ends at a very high point for Speaker An (Figure 7b) but not for Speaker Binh in Figures 7a and 8.

The rising part of *nga* varies from form to form (i.e., syllable to syllable, the non-tonal part) in the same speaker. In Figure 9 from Speaker Phuong, *nga* has a high ending in [ma] (9a) but a low one in [sa] (9b).

Figure 9a. *Nga* with a high ending in [ma] from Speaker Phuong
Figure 9b. Nga without a high ending in [sa] from Speaker Phuong

This tone also varies from token to token in the same speaker, as in Figure 10. Nga has a high ending in (10a), but the ending is low in (10b).
Figure 10a. *Nga* has a high ending in the first token of [ma] in Binh.
The data discussed in this section suggest that *hao* may or may not be curved (see Figures 7a and 7b) and, more importantly, the high ending in *nga* is not a necessary feature of the tone (see Figures 9b and 10b). The height of the ending varies from speaker to speaker and from token to token within a speaker. What appears to be consistent is that the tone *nga* always falls and rises to make a curve.

3.3.1.2.2. NANG1 and NANG 2

_Nang1_ is a falling tone with a glottal stop. This tone is said to end with a glottal stop (Vu 1988). However, what we usually see is that after the glottal stop or a very
creaky portion, the tone goes either flat or slightly down, or even sometimes slightly up (as in Figure 5d).

Unlike the descriptions in Vu 1999:40 and in Vuong and Hoang 1994, I found that nang2 does not fall steadily, but that it is just slightly higher than huyen, except that it is much shorter than huyen due to the final stop consonant (Figure 5h). This result agrees with that in Hoang 1986 (Figure 4).

3.3.2. Spectrographic analysis: waveform and spectrogram

In this section I first provide some background about different phonation types, namely creakiness and breathiness, and then show these qualities in Vietnamese tones using waveforms and spectrograms from my study.

3.3.2.1. Background in phonation type (modal, creaky, breathy)

In their study Nguyen and Edmondson 1987 use the term 'voice quality' to refer to phenomena such as pharyngealization, laryngealization, glottal stop, breathy voice and so on. Following Ladefoged 1988, I avoid the term 'voice quality' because in some work (Laver 1984) it also refers to the presence or absence of nasalization. As Ladefoged states, 'phonation type' is hard to define precisely if we want to include not only those states in which the glottis is vibrating in some way but also the use of the abducted vocal folds in voiceless sounds (1988:374). However this term is useful, and I adopt it here. Following Laver (1984:184) I assume that phonation is the action of the larynx to
transform the airflow into the audible acoustic energy that is necessary to perceive voice quality. Thus the term includes modal, breathy and creaky voice.

Figure 11 represents the waveforms of three laryngeal settings, modal, creaky and breathy, taken from Ladefoged and Maddieson 1996.

Figure 11. Waveforms of modal, creaky and breathy vowels spoken by one speaker of Jalapa Mazatec.

In Figure 11, modal voice is shown in the middle row. Modal voice is a laryngeal setting in which the arytenoid cartilages are in a neutral position for voicing, neither pulled apart nor pushed together (Stevens 1988, from Ladefoged and Maddieson 1996:50). This
modal voicing produces periodic pulses with a moderate level of frequency and intensity, characterized by regular cycles in the waveform in Figure 11.

Creaky voice is represented in the top row in Figure 11. In creaky voice, the vocal folds are strongly adducted and longitudinal tension is weak (Marasek 1997). This thickens the vocal folds and results in more irregular glottal pulses (more jitter) at a very low frequency (25 - 50 Hz). The amplitude is reduced a little bit more than that in modal voice.

Breathy voice is represented in the bottom row in Figure 11. In breathy voice, the vocal folds are weakly adducted, i.e., the contact between vocal folds is poor. Vocal fold vibration is not efficient and is at a lower frequency than in modal voice. Closure of the glottis is not complete. This results in glottal leakage, causing audible friction noise (Marasek 1997). In some cases it is very difficult to identify individual pulses in the waveforms, as we see in Figure 11. The glottal pulses are regular but the amplitude is noticeably reduced.

The technique I used to measure phonation types is the digital sound spectrography, discussed in Ladefoged, Maddieson and Jackson (1988). This is the method Kirk, Ladefoged, P. and Ladefoged, J. (1984) used to measure phonation types in Jalapa Mazatec (as reported in Ladefoged, Maddieson and Jackson (1988:298)). As Ladefoged, Maddieson and Jackson said, ‘this technique provides an excellent general view of many aspects of differences in phonation types (1988:300). These authors also investigated phonation types in Jalapa Mazatec by measuring the variation in the interval between glottal pulses in waveforms (1988:302) (see Dilley 1996 for more on this method).
However, it was not easy to replicate the spectral sections in Ladefoged and Maddieson 1996 with Vietnamese vowels due to technical problems. For the purpose of my thesis, a visual inspection of specific cues of phonation in the waveforms and wide-band spectrogram is sufficient. Thus in the next section, I show visual characteristics of these three phonation types in waveforms and spectrograms of Vietnamese vowels. Each pair of tones that shares the same contour but a different register is shown together, e.g., *ngang* with *huyen*, *sac1* with *nang1*.

3.3.2.2. Waveforms and spectrograms of the eight tones in Vietnamese

I show first the original waveforms of tones and their filtered counterparts in the same speaker, Son, followed by their spectrograms. The spectrographic data from other speakers are given in Appendix 6. The filtered waveforms help us identify individual pulses more clearly in breathy and creaky tones. All syllables have the same form [ta:] except for *sac2* and *nang2* with CVC syllables, exemplified here with the final consonant [k], i.e., [ta:k].

Figure 12a shows the waveform and filtered form of the two level tones, *ngang* and *huyen*. 
Figure 12a. Waveform of two level tones, *ngang* and *huyen* in [ta:] from Speaker Son.

The first row in (12a) is the original waveform of *ngang*, the high level tone. Its filtered counterpart is in the second row. This tone has a modal voice, as seen by the regular pulses and relatively high intensity (the peak of cycles is relatively high). There were the acoustic cues used to identify the presence of breathiness. In the third row is *huyen*, the low counterpart of *ngang*. This is a breathy tone characterized by periodic glottal pulses but with reduced amplitude after approximately 60 ms. This point can be seen clearly in its filtered counterpart in the last row of Figure 12a: it has very low intensity compared to that in *ngang*. Figure 12b shows spectrograms of these two tones.
It is more difficult to identify individual pulses in the spectrogram than in the waveform of the breathy *huyen* tone. Breathiness in *huyen* in (12b) is clearly indicated by irregularly space pulses and reduced amplitude, especially between 60ms and 100ms. This is the typical pattern for breathiness in this study.¹⁴ This characteristic of breathiness is not seen in the tone *ngang* where the pulses are periodic with moderate amplitude, represented by darker glottal pulses.

¹⁴ However, there is one female in my study whose breathy tone has a similar pattern to that in Figure 11, where it is impossible to see individual pulses.
Figures 13a and 13b show the waveforms and spectrograms of the rising tone sac1 and the falling tone nang1 in an open syllable, [ta:].

Figure 13a. Waveform of the rising sac1 and the falling tone nang1 in Speaker Son
In (13a) sac1 has modal voice with regular glottal pulses. Nang1 has creaky voice. However, in Vietnamese there are two manifestations of creaky voice: one has a glottal stop, and the other one is really creaky, similar to the pattern of irregular, low frequency pulses shown in Figure 11. The presence of either of these phenomena was used to identify creakiness in the soundwaves. Nang1 has a glottal stop in the middle of the tone, from approximately 100 to 200 ms, characterized by a very long non-periodic portion in the waveform, interrupted by two irregular pulses. In the spectrogram in (13b), the corresponding portion of this period is seen as irregular, widely space pulses. After this.
portion, the tone continues with a few irregular pulses. This is a typical pattern for *n邦* in various speakers.

Figures 14a and 14b show the waveforms and spectrograms of the two curved tones, *hoi* and *nga*.

Figure 14a. Waveform of the two curved tones, *hoi* with breathiness and *nga* with glottal stop in Speaker Son

As seen in the first row in 14a, *hoi* has breathy voice like *huyen* in Figure 12a. However, *hoi* is somewhat breathier than *huyen*. Its original (raw) waveform has weaker amplitude and it is more difficult to identify individual pulses in this tone. The reduced amplitude of *hoi* is clearly seen in the filtered form in the second row. This characteristic may be seen in the spectrogram in Figure 14b below in the gaps between pulses, from 40 ms to
60 ms. Especially between 70 and 120 ms, the tone is very breathy: glottal pulses are very weak, and it is difficult to see individual pulses.

Figure 14b. Spectrogram of *hoi* and *nga* in Speaker Son

In the other curved tone, *nga*, we see that the tone is broken in the middle with a glottal stop from 80 to 120 ms, interrupted by one or two widely space irregular pulses (second row in 14b). This corresponds to the wide gaps between the vertical striations in the spectrogram in the bottom row. After approximately 130ms, it has regular cycles again, although with lower intensity.
Figures 15a and 15b show the last two tones, sac2 and nang2, which occur in stop-final syllables. The rising tone sac2 has modal voice like sac1, but it is much shorter, ending after about 100 ms. The falling tone nang2 has some degree of breathiness compared to sac2: in the waveform (Figure 15a, bottom row), glottal pulses of nang2 are longer and weaker than those of sac2 (second row). Nang2 is also very short, ending after about 100 ms. However, unlike nang1, there is no glottal stop or creakiness in nang2.

Figure 15a. Waveform of the two tones sac2 and nang2 from Speaker Son
The visual inspection of the spectrographic characteristics of the eight Vietnamese tones is summarized below. Three tones have modal voice - *ngang*, *sac1* and *sac2*. Three tones have breathy voice - *huyen*, *hoi* and *nang2*. Finally, two tones show creaky voice - *nang1* and *nga*.

*ngang* modal voice (periodic, regular glottal pulses and moderate amplitude).

*huyen* breathy voice (regular glottal pulses, reduced amplitude).

*sac1* modal voice (regular glottal pulses and moderate amplitude).
**nang1**
glottal stop or creaky portion close to the end of the tone (irregular wide space pulses, sometimes interrupted by one or two irregular pulses; reduced amplitude. Complete closure of the vocal folds results in glottal stop; incomplete closure of the vocal folds results in creakiness).

**hoi**
breathy after approximately 40 ms, breathiest from 70 to 120 ms (regular pulses, reduced amplitude).

**nga**
glottal stop or creakiness in the middle of the tone (irregular wide space pulses from about 70 to 130 ms, periods of gaps in spectrogram).

**sac2**
modal voice (regular pulses, moderate amplitude), but very short due to the final stop, ending before 120 ms.

**nang2**
some breathiness after 60 ms (regular pulses, reduced amplitude), but also very short due to the final stop, ending before 120 ms. The F0 of nang2 is very close to that of huyen.

The characteristics of tones described above are quite consistent across speakers (see Appendix 6).

Although both hoi and huyen are breathy, hoi is breathier than huyen. However, breathiness in hoi decreases towards the end. Typically, in hoi the F0 of pitch rises after the breathy portion, giving the tone a curved shape. In the falling huyen tone, breathiness increases toward the end. Both nang and sac1 have modal voice. From a spectrogram it is not easy to tell the difference between these two tones. Their difference lies in the value of F0 of the ending: sac1 is a rising tone while nang is a level tone.
In the following section I briefly discuss the correlation between F0 and phonation types, the variation in degree of phonation types, and the results of a similar study in the literature on phonation types.

3.3.2.3. Discussion

There is a relationship between fundamental frequency and phonation type in Vietnamese. Tones with creaky or breathy voice also have low F0, either throughout the whole tone, e.g., huyen, nang2, or at some point, e.g., nang1, hoi, nga. This is not surprising, since the faster the vibration of the vocal folds, the higher the tone. In breathy voice, the vocal folds vibrate at a slower rate than in modal voice. This point will be discussed in Chapter Five, section 5.1.2.

Although breathiness and creakiness in huyen, hoi, nang1, nang2 and nga are found in almost every token, acoustic measurements often show some differences in degree. For instance, in the same speaker, although both huyen and nang2 are breathy, nang2 is less breathy than huyen. This is related to the shortness of nang2. Shortness prevents breathiness from being fully realized. Shortness does not cause breathiness, because sac2 is also very short but has modal voice. Usually in the spectrogram, phonation type is observed after about 60 ms. In nang2, the tone is very short. Breathiness is not very strong before the tone ends at around 110 ms; therefore, there is not enough time for breathiness to be fully realized. Moreover, with sac2 and nang2, two tones that occur in stop-final syllables, the primary cue to perception is not breathiness (there is no breathiness in sac2), but the complete blockage of the air flow due to the sudden closure of the oral cavity at the end of the tone.
In the Vietnamese literature, NE 1997 is perhaps the only study on Vietnamese tones that uses modern instrumental techniques to investigate phonation types. Using a Rothenberg mask to record the data and a method called 'Digital Inverse Filtering' (NE 1997:3), NE observed the size and shape of waveforms when airflow goes through the vocal folds. By looking at glottograms (a filtered waveform which tries to show only glottal vibration) from three speakers, NE found a large difference between the waveforms of six tones in terms of the adjustment of the vocal folds. Three tones, ngang, huyen and sac1, are relatively stable whereas in producing nang1, hoi and nga there are noticeable changes in the quality and quantity of the airflow shown in the glottogram.

NE found that in general, ngang has modal voice. Huyen shows breathiness for some speakers but not for all. Sac1 has modal voice. In one speaker NE describe a 'harsh' voice in hoi due to the vibration of both the vocal folds and the false vocal folds as well; it is not clear how NE define 'harsh voice'. Nga starts with modal voice, then in the glottogram the airflow is very small in the middle (there is little or no voicing due to creakiness or glottal stop), and ends with modal voice, so the waveform, they observe, looks like an hour glass, (horizontally). Nang1 also starts as modal voice, then the amount of air going through the vocal folds decreases towards the end.

These descriptions are largely in agreement with my results. The small difference is that NE claim that breathiness is optional in the tone huyen in some speakers. In my study, there is a difference in the spectrograms of ngang and huyen in all speakers, with huyen always being breathy and ngang modal. The degree of breathiness varies among speakers, as we will see in Chapter Five, 5.1.2. Note that NE use only glottograms (filtered waveforms) to compare speakers. I use both waveforms and spectrograms to
show the differences in phonation of tones in the same speaker. Because there is variation in the degree of breathiness or creakiness between speakers, I do not compare phonation types across speakers. It is more important that, like the difference of pitch range individually, a speaker uses phonation contrastively, given their own physical characteristics, e.g., one speaker might have a creakier voice than another.

However, although NE confirm that the phonation types of creakiness and breathiness are distinctive in Vietnamese tones, they do not attempt to show how these voice qualities are represented in the phonology, which I discuss in Chapter Five.

3.3.2.4. Summary

In summary, with waveforms and spectral information, I have shown in this section that among the eight tones in Vietnamese, ngang, sac1 and sac2 have modal voice, huyen, hoi and nang2 have breathy voice, and nang1 and nga have creakiness. These voice qualities are found consistently for all speakers. The two tones sac2 and nang2 are very short. Putting together the contour feature from Table 2, I summarize the phonation types in Table 3 as follows (for the moment I use the term register for phonation type. The motivation for this use is discussed and justified in Chapter Five). Table 3 shows the characteristics of features that remain stable in each tone. The parentheses around the feature (curve) in hoi shows that this feature might vary from speaker to speaker, as discussed in section 3.3.1.
Table 3: Stable features in each tone

<table>
<thead>
<tr>
<th>Tone</th>
<th>ngang</th>
<th>huyen</th>
<th>sac1</th>
<th>nang1</th>
<th>hoi</th>
<th>nga</th>
<th>sac2</th>
<th>nang2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contour</td>
<td>level</td>
<td>level</td>
<td>rise</td>
<td>fall</td>
<td>(curve)</td>
<td>curve</td>
<td>rise</td>
<td>fall</td>
</tr>
<tr>
<td>Register</td>
<td>modal</td>
<td>breathy</td>
<td>modal</td>
<td>creaky</td>
<td>breathy</td>
<td>creaky</td>
<td>modal</td>
<td>breathy</td>
</tr>
</tbody>
</table>

3.3.3. Other features: F0, length and linear portion of tone

In this section I present my findings concerning F0, length and linear portion (portion that contains critical information) of tones in Vietnamese. I demonstrate that F0 is not stable, tonal length is not distinctive, and that with complex contour patterns, the linear portion of Vietnamese tones is not fixed. It can be the midpoint or the end point depending on which tone one examines.

3.3.3.1. There is no difference in the F0 of tone

In this section I provide evidence that, in many cases, the height of F0 between two different tones is not different. Note that in this section, a tone in open syllables is sometimes compared with a tone in stop-final syllables, because I only look at the F0 difference between tones.

In the traditional literature on tonal pitch, *huyen* is classified as a low-level tone, and *nang2* as a low falling tone. Figures 16 shows the fundamental frequency of these two tones from Dung, a female speaker. The F0 difference between *huyen* and *nang2* is very small in (16b), and almost no F0 difference exists in (16a), where the two tones lie on top of each other. In Figures 16a and 16b, *huyen* is a long tone, *nang2* is a short one which ends with the final stop consonants.
Figure 16a. No F0 difference between *huyen* in [ka] and *nang* in [ka:ɔ] from Speaker Dung
Figure 16b. No F0 difference between *huyen* in [ta] and *nang2* in [ta:k] from Speaker Dung

Similarly, there is little difference in tonal height between the breathy curved *hoi* and the low falling *nang2* as in Figure 17. Recall that *huyen*, *hoi* and *nang2* all have breathy voice. As discussed in section 3.3.2.1, vibration of the vocal folds is at a lower frequency for breathy voice than for modal voice. This might explain why we sometimes find patterns such as those in Figures 16 and 17 where all tones are breathy, thus there is very little F0 difference between them.
Since both *huyen* and *nang2* are classified as low register, the fact that they are similar in F0 is not surprising. The similarity of *hoi* and *nang2* is more surprising, given that they in different registers phonologically (see Chapter Two). Even more surprising, *sac1*, traditionally a 'high' tone and *nang2*, traditionally a 'low' tone are almost identical in F0. Figure 18 shows the F0 of the high rising tone *sac1* and the low falling tone from Speakers Hoang.
In Figure 18 *sac1* is the long tone while *nang2* is the short one. F0 in these cases does not give any information for distinguishing the tones.

The above cases show that the height of F0 is not always the reliable cue that it is traditionally thought to be.

3.3.3.2. Variations in tonal length

In a number of studies, length is used (among other features) to describe tones phonetically (e.g., Doan 1977, Vu 1999), and even phonologically. Alves (1997) posits
[length] as an inherent feature of tone, based largely on his impressionistic perception. He suggests that *sac* and *nang* (Alves adopts the six-tone view, so *sac* and *nang* in stop-final syllable are the same as in nonstop final-syllable) are inherently short, *hoi* and *nga* are inherently long, and *ngang* and *huyen* are neither short nor long inherently. However, Figure 1 from NE 1997 shows that *sac* is longer than *nga*, and in Figure 2 from NE, we see that in speaker 2 (2b), *sac* is longer than *nga* and *ngang* is longer than *huyen*, but the opposite is true in speaker 6 (2a); *hoi* is short in speaker 6 (2b) but long in speaker 2 (2b). My study shows even more variation.

The two tones that are always short are *sac2* and *nang2*. Their shortness is likely related to the final voiceless stop. There are two important points to note: the shortness of *nang1* in open syllables, and the relationship between tonal length and vowel length. I will first discuss *nang1* in open syllables.

In the traditional literature on tones, many studies (Doan 1977:116, Vuong & Hoang 1994:98, Vu 1999:42 among others) claim that *nang1* ends with a glottal stop which makes the tone very short, even though this tone occurs in open syllables, e.g., /ma/ 'rice seedlings', /ta/ 'a hundred kilograms', /la/ 'strange'.

Therefore, short tones include *nang1*, *sac2* and *nang2* in their classification (Vu 1999:46). For example, Figure 19 from Speaker An shows the tone *nang1* ending with a glottal stop. *Nang1* is the falling tone in Figure 19. In this case, the tone drops abruptly and stops within 30 ms, at 120 ms.

---

15 These studies usually assume a six-tone hypothesis in which *nang1* and *nang2* are the same. A careful look at the phonetics of *nang1* (in open syllables) and *nang2* (in closed syllables) from my study shows that a glottal stop occurs in open syllables (Fig 13b) but not in closed syllables (Fig 15b). This fact is also observed in Cao 1988 and Vu 1982.
Figure 19. Nang1 ends with a glottal stop from Speaker An

Recall that except for sac2 and nang2, all tones can occur in sonorant-final syllable, which have equal length. (In Vietnamese it is said that the length if the rhyme in open syllable or sonorant-final syllables is always the same. This point is discussed in Chapter Six.) The unusual shortness of the tone nang1 has puzzled researchers. Some argue for phonemic status for the shortness of this tone to account for this fact (Vu 1999). However, my research shows that nang1 is not always short due to a glottal stop, as is claimed in the literature.

As described earlier in this chapter, first, the glottal stop occurs in nang1 (Figure 13b) but not in nang2 (Figure 15b) and sometimes there is no glottal stop, but a creaky
portion instead. If creakiness occurs, the tone is longer than if it ends in a glottal stop, because the vocal folds still vibrate during creakiness. Second, it is unusual to end this tone after the glottal stop; more often, after the glottal stop the tone nang1 either remains level or falls, or even rises slightly. Thus this tone is as long as other tones. 

Figure 20 shows six tones in open syllables from Speaker Binh. In this figure nang1, represented by a line with crosses, is as long as other tones.

Figure 20. Nang1 is as long as other tones

Figure 21 below shows nang1 with both short and long versions from the same speaker An. It is short in [ma] (21a) and [ka] in Figure 19, but long in [sa] (21b).
Figure 21a. The falling tone *nang1* is short in [ma] from Speaker, An.
Therefore, the shortness of nang1 is optional and varies from form to form, and from token to token. The shortness does not have distinctive status.

I now turn to tonal length with regard to vowel length. It is generally assumed that since tone is realized on a vowel, it is a property of a vowel. The phonological domain of tone is discussed in Chapter Six. Here I discuss only the phonetic domain of tone. So far we have focused on open syllables. Let us look at closed syllables with non-stop final consonants, i.e., syllables that end with either a glide or a nasal consonant. In these types of syllables, we see that the tone is spread over the whole rhyme, not just the
vowel. Let us use *nga*, the tone with most complex contour (it is creaky or has a glottal stop in the middle of the tone), to see how the whole contour can be realized. Figure 22 shows the tone *nga* on a form with a vowel and a final nasal consonant, [taːm]. As we might expect, the creaky portion of glottal stop occurs approximately in the middle of the rhyme, that is between the end of the vowel and the beginning of the nasal (from 90 ms to 150 ms in the waveform). The rising part is realized on the nasal. Note that although we find only extremely weak voicing in the nasal portion (after 150 ms), it is still essential for the tone to be fully realized. If the nasal part is cut off, the rising part is missing. As a result, the tone becomes unrecognizable.

Figure 22. Tone *nga* in a nasal-final syllable with a long vowel [taːm]
If the vowel is short as in [tam] in Figure 23, we still find the same phenomenon except that now the vowel part is extremely short (it ends before 60 ms); the glottal stop occurs instead of the creaky portion and it is quite long (from approximately from 60 to 120 ms); and the nasal part is lengthened (starting after 120 ms, compared to 150 ms in [tam] in Figure 22). More interestingly, the length is almost the same in both Figures 22 and 23 (210 ms). This fact tells us that vowel length does not affect syllable length.

Figure 23. Tone nga in a nasal-final syllable with a short vowel [tam]

In summary, I have shown that tonal length can vary from token to token and from speaker to speaker. It can be distributed either over the entire vowel in open syllables, or over the whole rhyme in non-obstruent final syllables. Tonal length, therefore, is not distinctive.
3.3.3.3. Tone linear portion

The last point I want to discuss here is the linear portion of tones, i.e., the portion that contains the most important information of a tone. Is it the onset, midpoint or endpoint of the tone that is important? I will show that in Vietnamese, the onset is not significant. In non-curved tones, e.g., ngang, huyen, sac1, sac2, nang1, and nang2 the crucial information usually comes late in the tone, and sometimes very late, during about the last 30 ms.

If we look at Figures 24 and 25 from the two males, the onsets of tones are not very different. It is even difficult to separate the onsets of some tones. Note that in these figures, the y-axis scale is different from that in females. For males, the pitch range is scaled maximally (180 Hz, compared to 350 Hz in females). Figure 24a shows the 8 tones in [ta]/[ta:k] from Speaker Hoang.
Figure 24a. Eight tones in [ta] and [ta:k] from Hoang

Figure 24b shows only the first 60ms of each tone from Figure 24a.
Figure 24b. The first 60 ms of 8 tones in [ta]/[ta:k] from Hoang

In this figure, it is hard to see anything distinct among tones, except that nga goes down noticeably. The fact that the first 60 ms does not show the difference in F0 between tones perhaps to prevent tone from being affected by the neighboring tone.
Figure 24c shows the same eight tones from Hoang, omitting the first 60 ms.

Figure 24c. Eight tones in [ta]/[ta:k] from Hoang omitting the first 60 ms

This figure shows all tones distinctively, especially tones which rise or fall immediately after the first 60 ms. Nang2 does not fall steadily because, like huyen, it is breathy (see more in chapter Five).

Figure 25 shows eight tones in Son, another male speaker.
Figures 24 and 25 show that in general, tones of the two speakers Hoang and Son have similar shapes except for hoi and nang1. Hoi goes down in Figure 24a, but not in Figure 25. Nang1 just goes down in Figure 24a but it goes down and then rises up in Figure 25. Furthermore, nga has glottal stop in Figure 25 but not in Figure 24a.

There are many cases where the fundamental frequency of the same tone randomly varies from token to token. For example in Figure 24a, the two level tones ngang and huyen have a steady fundamental frequency through out; nga falls abruptly after the onset. With other tones, only after about the first 60 ms, the difference in F0 becomes very noticeable, e.g., the fall of hoi or the rise of nang1 in Figure 24c). This is a
consistent finding with the patterns in spectrograms. After about the first 60 to 70 ms, the laryngeal features of creakiness and breathiness become obvious. For instance, when we look at the spectrogram of huyen in Figure 12b, the pattern of breathiness in the waveform and spectrogram starts before about 60 ms. In Figure 13b, the glottal stop in nang1 starts at 90 ms.

Now let us see where crucial cues in a tone appear. In the two level tones, ngang and huyen, it is obvious that the information is given throughout the whole tone after approximately the first 60 ms. The phonation type that the tone should have, e.g., breathiness in huyen, is consistent throughout the whole tone. For the other tones, the essential information is given either during the endpoint or in the middle of the tone depending on what feature the tone in question has.

Let us first look at the rising tone sac2 and the falling tone nang2 (Figure 15b). These two tones are very short, as we have seen, approximately 90 to 100 ms in duration. The most important cue in these tones is that they end very suddenly, as represented segmentally by a final stop consonant. To distinguish between the two, nang2 has breathiness, here seen as longer and weaker pulses. Breathiness increases toward the end. The crucial portion of these two tones, therefore, is the endpoint.

With the rising and falling tones in non-stop final syllables, sac1 and nang1, the essential information comes in the second part of the tone. Usually sac1 does not rise until it gets to the second part of the tone, e.g., with Speaker Binh (Figure 20) or in most speakers in Figure 5c, where the tone does not rise until after 120 ms. However, it does not rise until two-thirds of the way into the tone with Speaker Hoang (Figure 18), or even until the last 30 ms in An (Figure 21). With the falling tone nang1, the glottal stop or
creakiness comes in the second part of the tone. Although usually the tone goes lower before it ends, the most important cue is the glottal stop. That is why we see that sometimes the tone is very long as in Figures 20 and 21b, and sometimes it is short as in Figures 19 and 21a. Most of the time, the portion after the glottal stop is level as in [ka] in Figure 21b, or occasionally it rises as in [ta] from Speaker Son in Figure 25. However, it is much more common that nang1 flattens or drops slightly after the glottal stop, and the rising part of nang1 is optional.

With the two curved tones, hoi and nga, as the tonal shape suggests, the most crucial part is the middle of the tone (the curve) especially for tone nga. Recall that hoi has breathiness and nga has either a glottal stop or creakiness. After the glottal stop or creakiness in the middle, nga rises. Here the rising part after the glottal stop is distinctive. Without this rising part, nga might be recognized as nang1, which also has a glottal stop. For instance, in Figure 24c where the first 60 ms is cut off, the creaky portion in nga is shown by the fact that the tone has a lowest starting point there.

With the breathy curved tone hoi, in most speakers, the curve is always present, although how curved it is varies from speaker to speaker and form to form (e.g., Figures 7a and 8 from Binh, 5e from eight speakers). The fact that the curve in nga is much more obvious than the curve in hoi has a physiological explanation: during the glottal stop, subglottal air pressure builds. After the vocal folds open, the air is suddenly released, causing a burst of high frequency vibration. Hoang Cao Cuong (1986:27) also comes to the same conclusion based on his experimental work, that the first part of a Vietnamese tone gives less information than the second part. However, he does not distinguish among tones with different contours.
In summary, I have shown that the linear portion in Vietnamese tones is the midpoint in curved tones and the endpoint in non-curved tones.

3.4. Summary

In this chapter, I have presented the findings from my experimental study, using pitch graphs and spectrograms. I have shown that the phonation types of breathiness and creakiness are found in huyen, hoi, nang1, nang2 and nga. More importantly, in many cases, I have found that F0 of pitch variable within and across speakers. It does not give information that allows one to distinguish tones when two different tones are compared. Moreover, there is a correlation between phonation types and F0: where we find breathiness or creakiness, the tone is low, either in the entire tone if breathiness or creakiness occurs throughout, or at the point where they are found. I have also shown that tonal length and height can vary from speaker to speaker, form to form, and token to token. A follow-up perception test would tell us whether without phonation types, tones are not recognizable in Vietnamese.

I call a phonetic feature that invariant across and within speakers a primary feature. If the F0 of pitch is not a primary feature of Vietnamese tones, there must be something else that is stable, and hence more reliable. I will return to this issue in Chapter Five. First, however, I address an issue that I put aside in Chapter Two, namely the number of tones in the system.
The size of the Vietnamese tonal inventory is a topic of much debate in the literature, both in terms of the number of tones and the number of features required to classify them. I have so far focused largely on the tones that occur in open and sonorant-final syllables, ignoring the two tones that occur in stop-final syllables. Two different positions are taken with respect to these latter tones. The six-tone hypothesis argues that these tones belong with sac and nang; the eight-tone hypothesis treats them as independent tones. In this chapter I first present different views of the tonal inventory in the literature, and then I argue for the eight-tone system. The evidence for this hypothesis is drawn mainly from phonology, with some acoustic evidence in addition. I show that sac2 and nang2, the two tones that occur in stop-final syllables, are different from those in non-stop final syllables. I also provide formal representations for these two tones and argue that sac2 is less marked than nang2.

4.1. The tonal inventory: are there six or eight tones?

In general, there are two major views concerning the size of the underlying tonal inventory of Vietnamese: (i) Vietnamese has a six-tone system with both nasals and stops being phonemic in final position (Doan 1977, Vuong & Hoang 1994 and others); and (ii) Vietnamese has an eight-tone system with no manner distinction in final
consonants (Cao 1998, Hoang Thi Chau 1989 and the view assumed in this paper). Note that smaller inventories are also proposed (see Doan 1977, Vu 1988).

In this section, I examine each of these views. I first briefly discuss the phonetic differences between $sac_1$ and $sac_2$, and between $nang_1$ and $nang_2$. $Sac_1$ and $sac_2$ are rising tones, $nang_1$ and $nang_2$ are falling tones. Recall from Chapter Three, section 3.3, that, in terms of length, $sac_2$ and $nang_2$ are very short compared to their counterparts $sac_1$ and $nang_1$; in terms of phonation types, $nang_1$ has either glottal stop or creakiness while $nang_2$ has some breathiness instead.

4.1.1. A six-tone system

In Vietnamese, a limited number of segments can occur in final position: glides, nasals and voiceless stops (see Chapter One, section 1.2.3). In non-stop final syllables six tones are found, while in stop-final syllables only $sac_2$ and $nang_2$ occur. The tonal distribution given in Chapter Two is repeated here as (49).

(49) Tonal distribution

<table>
<thead>
<tr>
<th>Syllable types</th>
<th>e.g.</th>
<th>ngang</th>
<th>huyen</th>
<th>sac1</th>
<th>nang1</th>
<th>hoi</th>
<th>nga</th>
<th>sac2</th>
<th>nang2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV</td>
<td>ba</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVG</td>
<td>ba:w</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVN</td>
<td>lyn</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVT</td>
<td>ba:t</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

138
Recall also that only six tones are distinguished in the orthography of Vietnamese: ngang, huyen, hoi, nga, sac and nang (see Chapter 1, section 2.1.1). Therefore, the orthography assumes that sac and nang occur in both syllable types while other tones only occur in open or sonorant-final syllables. The discrepancy in distribution between sac and nang and the other tones leads to an interesting question. Assuming six tones, only sac and nang appear in all syllable types, as seen in (1), while the others are restricted to syllable types that are not closed by obstruents.

The standard position in the Vietnamese literature, reflected in the orthography, is that there are six tones in the inventory with final consonants distinctive in manner (Doan 1977, Vuong and Hoang 1994 and others). In this view, sac1 and sac2 and nang1 and nang2 are allophones, being in complementary distribution, i.e., sac1 and nang1 occur in non stop-final syllables while sac2 and nang2 occur in stop-final syllables. There is no phonological tonal distinction between sac1 and sac2 or between nang1 and nang2 under this hypothesis. The distinction lies rather with the final stop or sonorant, which means tone is predicted by manner of the consonant. For instance, sac is realized as sac1 in an open syllable or before a sonorant, but as sac2 before an obstruent. The other tones have only a single form, as expected, as no manner contrast is present in consonants following them. Thus, consonantal manner is basic and predicts tonal shape. The phonological tonal inventory of a six-tone system, based on tonal patterning in reduplication, is shown in (50), where the sacs and nangs are categorized together, i.e., non-even, non-broken. I assume that the two registers are unmarked and marked (see Chapter Two, section 2.2.1).
Let us set aside \textit{sac} and \textit{nang} for a moment and consider manner of articulation of final consonants. With the other tones, this can be considered to be redundant: only sonorants follow \textit{ngang}, \textit{huyen}, \textit{hoi} and \textit{nga}. This predictability of manner disappears when \textit{sac} and \textit{nang} are considered, however: both sonorants and obstruents follow \textit{sac} and \textit{nang}; therefore, with these tones, manner must be distinctive. In these cases, manner of articulation must be listed as part of the lexical entry. An important question we can ask then is why final stops are so restricted in their distribution. This question has been of some concern in the traditional literature on Vietnamese tone. Doan 1977 and Hoang 1986, among others, offer an explanation based on phonetics. According to them, a rhyme with a final voiceless stop is too short for other tones to be realized, e.g., \textit{ngang} and \textit{huyen} need to occur on a long enough span to show their level contour, and \textit{hoi} and \textit{nga} need an appropriate length to show their complex contour. \textit{Sac} and \textit{nang} are simply rising and falling, their contour is neutralized and they only need to be distinguished by their registers (Hoang 1989). However, this explanation based on timing is not satisfactory across dialects and languages for several reasons. First, in the Hue and Saigon dialects a phonetically complex contour tone can occur with the final stop (Hoang Thi Chau 1989, Vu 1982, Hoang Cao Cuong 1989). For instance, Figure 26 from Hoang Cao Cuong 1989 shows the tonal system in the Hue dialect.
Figure 26. The phonetically complex contour of sacs in the Hue dialect (Hoang Cuong 1989)

This figure shows the tonal contour of seven tones. Note that in the Hue dialect, nga is neutralized with hoi. In this figure only sac1 and sac2, the two tones under
consideration, are labelled. Speaker 1 is male, speakers 2 and 3 are female. Hoang does not label duration of tones in the diagram. Examining the clearest pitch graph, that of Speaker 2, we see that the two *sacs*, which are rising tones in the Hanoi dialect, are falling-rising tones in Hue. They are in fact the only curved tones in Hue. Therefore, in the Hue dialect, the rising tones *sac1* and *sac2* in the Hanoi dialect surface as the curved tone *hoi*, i.e., a falling-rising tone, in the Hanoi dialect.

Figure 27 is also taken from Hoang Cao Cuong 1989. This figure shows the pitch graph of four informants from the Saigon dialect. Speakers 1 and 2 are male, speakers 3 and 4 are female.
Figure 27. The phonetic complex contour of *nangs* in the Saigon dialect (Hoang 1989)

In this figure, I label only *nang1* and *nang2*, as these are our focus. These two tones, which are falling tones in the Hanoi dialect, are falling-rising tones in the Saigon dialect.

The third curved tone in this dialect is *hoi*. Recall that the absence of *hoi* and *nga* in stop-final syllables in the Hanoi dialect has been attributed to phonetics: the rhyme is too short to carry the complex contour tones. Figures 26 and 27 demonstrate that this proposal cannot explain why a tone with the same phonetically complex contour can
surface with stop-final syllables in the Hue and Saigon dialects. A second reason that the phonetic account is unsatisfactory has to do with implication. In other tone languages, the presence of a rising or falling tone in the system implies the existence of a level tone in that system (e.g., Anderson 1978:151). Because of this implication, researchers have tried to answer the question of whether a contour tone should be treated as a single unit or a sequence of two level tones (Anderson 1978, Yip 1995). Woo 1969 claims that the universal system of tonal features includes only a level tone element (Anderson 1978). If stop-final syllables are an environment that neutralizes tonal contours, one might expect that the two simplest tones, ngang and huyen, should occur in that environment. In fact, there are languages where stop-final syllables only allow level tones, such Be (Hashimoto 1985), a Be-Tai language in the Tai-Kadai family, or Sgaw and Pwo (Benedict 1972), two Karen languages in the Sino-Tibetan family. Both level and contour tones in non-stop final syllables are found in these languages. Yip (1995:487) reports that Cantonese, a Chinese language, has seven distinctive tones in non-stop final syllables, but only three level tones in stop-final syllables.

Burton 1992 tries to answer the question of why only sac2 and nang2 can occur in stop-final syllables. Burton's classification of tones in the Hanoi dialect, given in Chapter Two, section 2.4.1, is repeated as (51).

(51) Representation of Vietnamese tones (Burton 1992)

<table>
<thead>
<tr>
<th>REGISTER</th>
<th>CONTOUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ HIGH</td>
<td>A = 1 (ngang)</td>
</tr>
<tr>
<td>- HIGH</td>
<td>D = 1 (huyen)</td>
</tr>
</tbody>
</table>
Burton argues that only *sac* and *nang* can occur in stop-final syllables because they have simple contours, represented as 'h' in (51). *Hoi* and *nga* cannot occur in stop-final syllables because they have complex contours, represented as 'lh' in (51). Then an important question we might ask is why *ngang* and *huyen*, which also have simple contours, represented as 'l' in (51), cannot occur in stop-final syllables. In order to answer this question, Burton proposes a constraint that 'l'-pitch tones cannot occur before a voiceless consonant'. This, he argues, has phonetic motivation: there is a relation between voicelessness and high pitch (i.e., initial voiced consonants raise the pitch of the vowel, voiceless consonants lower the pitch (Ohala 1973)). Recall from Chapter One, section 1.2.3 that all final stops in Vietnamese are voiceless. *Ngang* and *huyen* are 'l', therefore, they cannot occur with final stops, which are voiceless. *Sac* and *nang* are 'h' tones; it follows that they can occur before voiceless consonants. He does not provide further explanation for this constraint.

Although this constraint follows from the affinity between low tone and voicing, there is some problem with this argument. Let us see how Burton classifies tonal contours based on phonetics. Figure 28 shows a pitch graph of the Hanoi dialect from Vu 1982. In this figure, F0 is in FD percent plotted against normalized duration.\(^\text{16}\)

\(^{16}\)To compare the F0 data, Vu devised a method of normalization which involves the notion of F0 Differential as a function (FD) of the mean F. The formula is given as follows:

\[
FD = \text{Itg} \left( \frac{F_{\text{avg}} \times 100}{F} \right)
\]

where \(F_{\text{avg}}\) is any individual F0 value, \(F\) is the mean F0 of a sample, and Itg is an integer (1982:68).
Figure 28. Normalized Mean F0 in the Hanoi dialect (Vu 1982)

Burton argues that *ngang* and *huyen* are classified as 'I' because they are articulated in the lower part of their respective registers. Because Burton bases his arguments on the acoustic study by Vu 1982, I assume that according to Burton, the middle line in Figure 28 divides tones into two registers. Thus, Figure 28 shows that *ngang* and *huyen* are articulated in the lower part of their register (*ngang* is in the upper register and *huyen* is in the lower register), therefore, they are 'I'. However, this argument does not hold with
nang, which is classified as 'h'. From Figure 28, one can assume that the sacs (sac1 and sac2) occur in the higher part of their register, therefore, they are 'h' tones. The problem arises when we look at nang1 and nang2. Except for the beginning part, both nang1 and nang2 are also low in their register, even lower than huyen. Recall from Chapter Three, section 3.3.3, that the endpoint, not the beginning, of nang1 and nang2 is distinctive. With this reasoning, it is difficult to maintain that the nangs are 'h' based on their phonetic properties. In addition, recall that Burton says that sac and nang can occur with final stops because they are high in pitch height. Recall also that final stops are always voiceless. Because voiced final stops do not exist in Vietnamese, it is not possible to test whether sac and nang could occur with voiced stop consonants.

In summary, the six-tone hypothesis, represented in the orthography, assumes a six-tone inventory. Manner of articulation of final consonants is partially predicted by tone: ngang, huyen, hoi and nga occur only with non-stop final consonants and in open syllables. Sac and nang can occur with stop or non-stop final consonants, with different allophones in each environment.

4.1.2. An eight-tone system

I now consider the second view, that there are eight tones in the phonological inventory (Doan 1977, Cao 1998, Hoang Thi Chau 1989).

Traditional poetry classifies tones according to their register and contour. There are two registers, 'phu' (high) and 'tram' (low). Contour tones are divided into 'binh' (even) and 'trac' (non-even) tones. Among non-even tones, there are 'thuong' (falling and rising) tones, 'khu' (rise or fall) and 'nhap' (checked) tones, i.e., the two tones that
occur in stop-final syllables. These terms were used as the names of the tones in the literature before the current terms were introduced (*ngang, huyen* and so on).

Note that in the system based on tonal patterning in poetry that uses the traditional names (e.g., *phu, tram* and so on), with respect to the two tones *hoi* and *nga*, there is also an inconsistency between researchers about the register of *hoi* and *nga*. While some (e.g., Hoang T Chau, 1989:202, Phan, 1997:23) classify *hoi* as *‘phu’* (high) and *nga* as *‘tram’* (low), others (e.g., Doan, 1977:120) put *hoi* in the *‘tram’* register and *nga* in the *‘phu’* register. The classification in (52) shows the former view. Note that phonetically, the traditional eight-tone view also classifies tones as the six-tone view does (see (50) where phonetic labels are not completely shown), with *hoi* being phonetically low and *nga* phonetically high.

(52) Classification of tones in traditional poetry

<table>
<thead>
<tr>
<th></th>
<th>trac</th>
<th>(non-even)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>binh</em> (even)</td>
<td><em>thuong</em> (fall-rise)</td>
</tr>
<tr>
<td><strong>PHU</strong> (high)</td>
<td><em>ngang</em></td>
<td><em>hoi</em></td>
</tr>
<tr>
<td><strong>TRAM</strong> (low)</td>
<td><em>huyen</em></td>
<td><em>nga</em></td>
</tr>
</tbody>
</table>

Phonologically, rather than being predictable variants of *sac* and *nang* in nonstop-final syllables, *sac1* and *nang1* are phonologically distinct from the tones in stop-final syllables, i.e., *sac2* and *nang2*. This hypothesis claims that the contrast between *sac* and *nang* in stop-final syllables and *sac* and *nang* in sonorant-final syllables lies with the tones, and the manner of the final consonant is conditioned by tone. For instance, *[mat]* with the tone *sac2* and *[man]* with the tone *sac1* contrast in the two different *sacs*, not in
the two different final consonants. Thus, there is no underlying manner contrast in final consonants. Because voiceless stops can only occur with sac2 and nang2, voiceless stops are conditioned by these tones: sac2 and nang2 predict voiceless stops. It is thus manner which is allophonic rather than realizations of sac and nang.

An underlying eight-tone system, together with an example, based on Cao 1998, is shown in (53). The example in (53), aC, shows Vietnamese orthography for tones in which the final consonant is unspecified for manner, represented by ‘C’. (Cao 1998 uses only the coronal in the example.) In the surface forms, ‘N’ represents a final sonorant. ‘T’ represents a final stop. Two registers, ‘phu’ (high) and ‘tram’ (low), have four high tones (ngang, nga, sac1, sac2) and four low tones (huyen, hoi, nang1, nang2). In terms of contour, the even tones are ngang and huyen. The six non-even tones are divided into three groups: curved tones (hoi, nga), rising/ falling tones (sac1 and nang1), and two additional tones (sac2 and nang2) which are non-curve. Note that like (50) and (52), I place hoi in the high register and nga in the low register, following the discussion in Chapter Three but contrary to Cao 1998.
(53) Underlying tonal inventory in an eight-tone system

<table>
<thead>
<tr>
<th></th>
<th>EVEN</th>
<th>NONEVEN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rising/falling</td>
<td>curve</td>
</tr>
<tr>
<td>Phu</td>
<td>aC (ngang) =&gt; [aN]</td>
<td>aC (sac1) =&gt; [aN]</td>
</tr>
<tr>
<td>Tram</td>
<td>âC (huyen) =&gt;[âN]</td>
<td>aC(nang1) =&gt;[aN]</td>
</tr>
</tbody>
</table>

Diacritics are orthographic representations of tone. The eight-tone view allows the manner of the final coronal to be predicted from tone. The difference between, for instance, [aN] (sac1) and [âT] (sac2) lies with tones. Sac1 and nang1 predict a final sonorant. Sac2 and nang2 predict a final stop. These are thus represented as /aT, sac1/ and /aT, sac2/ rather than /aT, sac/ and /aN, sac/ as in the six-tone view.

Cao 1998 captures this distribution by assuming that [±nasal] is a tonal feature which is realized on the final consonant. Sac2 and nang2 are [-nasal]; the final consonant, therefore, surfaces as a stop. Sac1, nang1 and the other remaining tones are [+nasal]; the final consonant surfaces as a sonorant. In (54), C is a consonant that has no manner specification.

(54)  /báːC/  (sac1 [+nasal])  [báːn]  to sell
      /báːC/  (sac2 [-nasal])  [báːt]  bowl
      /háːC/  (nang1 [+nasal])  [háːŋ]  rate, level
      /háːC/  (nang2 [-nasal])  [háːk]  crane
Cao comments that the writing system, by using the same diacritics for these two tones in all syllable types, is responsible for misleading phonologists to propose a six-tone inventory. If the founders of the current writing system had used distinct diacritics for sac and nang in stop-final syllables, this confusion of tones would not be a problem in the literature. For instance, in the table in (53) all consonants would share manner, but use two different diacritics for sac and nang in nonstop-final and stop-final syllables.

4.2. Advantages of the eight-tone view

In this section I examine several arguments for the eight-tone hypothesis. I will show that reduplication and the historical development of Vietnamese tones give support to the eight-tone inventory.

4.2.1. Support from traditional poetry and reduplication

First, while the eight-tone analysis is little discussed in the current literature, it is assumed in the presentation of traditional poetry before the current Vietnamese writing system had been created (Cao 1998) as discussed in section 4.1.2.17 Second, adding two tonemes reduces the phonemic final consonant inventory from six to three (i.e., from /p, t, k, m, n, ŋ/ to /M, N, ŋ/), where /M, N, ŋ/ represent places of articulation without specification for manner18, and gives a simpler account for tone harmony in reduplication (Cao 1998). For instance, the reduplication process

---

17 Cross registers, 'even' and 'non-even' are a basic division of tones in poetry. There are different patterns of alternation between even and non-even tonal groups in different verses, and even between tones of the same register in a rhyme position (see Doan 1977). Patterns of alternation are highly controlled by specific constraints.
discussed in section 2.3.2, Chapter Two, shows that the reduplicant tone must be either *ngang* or *huyen* depending on the register of the base tone. It is *ngang* if the base tone is an unmarked register tone and *huyen* if the base tone is a marked register tone. Bearing tone *ngang* or *huyen*, the reduplicant, therefore, must end in a sonorant consonant. If the base tone is other than *sac2* and *nang2*, as in (55), tone harmony is represented basically the same in the two views. Recall that the base tone is underlined.

(55) Tone harmony in reduplication if the base tone is other than *sac2* and *nang2*

```
/toj/ (sac1)  ‘dark’ > toj (ngang) + toj (sac1)   ‘rather dark’
```

Interestingly, if the base ends in a stop, the reduplicant must still have either *huyen* or *ngang* as its tone. These tones are incompatible with final stops, and a final nasal surfaces instead. This can be seen in (56), where the base tone *sac2* (56a) has the reduplicant tone *ngang*, which surfaces with a final velar nasal, or the base tone *nang2* (56b) has the reduplicant tone *huyen*, which surfaces with a final labial nasal.

(56)

a. *sac* (sac2)   ‘sharp’ > sang *sac* (ngang - sac2)   ‘rather sharp’

---

18 The issue of underspecification has never been discussed in the Vietnamese traditional literature; and in the eight-tone view, the final inventory is usually represented as underlying nasals as in (54) (e.g., Doan 1977:156, Vuong and Hoang 1994:98). However, by assuming that the underlying final consonants are nasals, these authors have both eight tones and unpredictable manner of the final. This thesis assumes the final consonant in Vietnamese is unspecified for manner. Glides can be treated as vowels (see discussion in section 4.3.1).
In this case, tone harmony is represented differently in the two hypotheses, as shown in (57). The six-tone analysis requires a rule to change a stop into a homorganic nasal, i.e., a final stop consonant must change to a homorganic nasal if it occurs with tones other than sac₂ and nang₂ (57a). The eight-tone hypothesis (57b), on the other hand, does not need this rule, as the manner of the consonant is predictable from the tone, i.e., ngang predicts a final sonorant. In the eight-tone hypothesis, I represent the final consonant, which is, recall, unmarked for manner, with a nasal.

(57) Tone harmony: sak ‘sharp’ (sac₂) > sanj sak (ngang - sac₂) ‘rather sharp’

a. Six-tone hypothesis

Tone harmony

/sak, sac/ ‘sharp’ > sak (ngang) + sak (sac) ‘rather sharp’

Nasalization rule for the reduplicant with tone ngang:

sak (ngang) > [sanj] (ngang)

b. Eight-tone hypothesis

Tone harmony and realization of the final consonant:

/sanj, sac₂ / ‘sharp’ > [sanj] (ngang) + [sanj] (sac₂) = sanj sak

The eight-tone hypothesis provides a more elegant account for this process of reduplication in Vietnamese as the manner is simply a consequence of default.
However, the eight-tone hypothesis comes at some cost, namely the enrichment of the representations by two tones, sac2 and nang2. Given my assumptions about markedness and phonological theory this hypothesis seems to be the simpler choice.\textsuperscript{19} I return to this discussion in section 4.3.

4.2.2. Historical development

Hoang Thi Chau (1989:202) supports the eight-tone view based not only on arguments from traditional poetry (shown in (52)), but also because the eight-tone view agrees with the hypothesis that tones were developed from segments. Ironically, because the historical hypothesis also assumes a six-tone system (see (58)) below), Hoang only adopts the idea from the historical hypothesis that, originating from segments, tones still have some relic of those segments as their properties. According to Haudricourt 1954, Vietnamese tones developed from voicing of the initial and the final segments, as shown in (58). Haudricourt uses p, b to represent initials - only the voicing is important here. s and x represent the voiceless fricatives, alveolar and velar, respectively. There are no final stops in (58) except ?.

\textsuperscript{19} Note that in terms of register preservation, like sac1 and nang1, sac2 and nang2 pattern with other unmarked and marked tones, respectively. For instance, in the following example, sac2 patterns with hoi, and nang2 pattern with huyen.

\begin{align*} 
  \text{a. } \text{vat (sac2)} & \quad \text{‘laborious’} & \quad \text{> vat va (sac2 - hoi)} & \quad \text{‘very hard’} \\
  \text{b. } \text{ngat (nang2)} & \quad \text{‘severe, hard’} & \quad \text{> ngat ngheo (nang2 - huyen)} & \quad \text{‘very hard’} 
\end{align*}
(58) Historical development of Vietnamese tones

<table>
<thead>
<tr>
<th>Original</th>
<th>Sixth century</th>
<th>Twelfth century</th>
<th>Modern</th>
</tr>
</thead>
<tbody>
<tr>
<td>pa</td>
<td>pa (ngang)</td>
<td>pa (ngang)</td>
<td>ba (ngang)</td>
</tr>
<tr>
<td>ba</td>
<td>ba (ngang)</td>
<td>pà (huyen)</td>
<td>bà (huyen)</td>
</tr>
<tr>
<td>pas / pah</td>
<td>pà (huyen)</td>
<td>pà (hoi)</td>
<td>bà (hoi)</td>
</tr>
<tr>
<td>bas / bah</td>
<td>bà (huyen)</td>
<td>pà (nga)</td>
<td>bā (nga)</td>
</tr>
<tr>
<td>pax / pa?</td>
<td>pà (sac)</td>
<td>pà (sac)</td>
<td>bā (sac)</td>
</tr>
<tr>
<td>bax / ba?</td>
<td>bà (sac)</td>
<td>pà (nang)</td>
<td>bā (nang)</td>
</tr>
</tbody>
</table>

Under this hypothesis, early Vietnamese was non-tonal. By the 6th century, three tones were established: *ngang* in open syllables, *huyen* from the final fricatives [s] and [h], and *sac* from the final [x] and glottal stop. By the 12th century, the voicing contrast was lost in the initial consonants, and gave rise to six tones (two registers). The original voiced initial gave rise to low tones (*huyen, nga, nang*). However, it is not clear how *hoi* and *nga*, for instance, developed from 'huyen'. Finally, in modern Vietnamese, a voicing contrast is reestablished and six tones remain.

It is claimed in the literature (e.g. Hoang Thi Chau 1989, Vu 1988) that when final consonants disappeared, they created tones with their trace: the final stops disappeared, leaving glottalization in tones *sac* and *nang*; final fricatives disappeared, leaving creakiness in tones *hoi* and *nga*. Therefore, establishing how tones developed also helps
to understand the origin of phonation types in the current tonal system. However, what the final *-x from *pax and *bax leaves in tone is not mentioned in this analysis.

This hypothesis is very influential in the Vietnamese literature. The historical development supports the eight-tone view in that it shows an intimate relationship between the final stops and tones.

Based on Haudricourt’s hypothesis, the mismatch between the phonetics and phonology of the two tones hoi and nga in reduplication (Chapter Three and Chapter Five) is explained in the traditional literature as switching of registers between these two tones, without questioning the validity either of the hypothesis or of the acoustic experiments. As shown in Chapter Three and discussed more in Chapter Five, it turns out that, with respect to hoi and nga, Haudricourt’s hypothesis still holds: there is no switching of registers between these two tones. Hoi patterns with ngang and sac, nga patterns with huyen and nang. My study, therefore, adds additional support to Haudricourt’s hypothesis from a different perspective.

However, there are some drawbacks to Haudricourt’s hypothesis. For instance, it is reasonable to claim that the historical final glottal stop left a trace in nang1 where the tone has either glottal stop or creakiness (see Chapter Three). However, it is hard to say the same for sac1, where glottal stop is not found in the tone, as pointed out in Alves 1997 (Alves also adopts the six-tone view). Similarly, hoi and nga developed from final *-s and *-h, therefore, it is said that the breathy voice found in these tones has its origin in these fricatives. However, as seen in Chapter Three, hoi has breathiness while nga is characterized by creakiness or glottal stop, which should come from the historical final glottal stop. Researchers try to explain the articulatory correlation between the final
consonants and phonation types found in tones, but it is not always clear. For instance, Vu 1988 suggests that breathiness from the historical fricative *-h might have changed to the phonation of glottal stop found in the current nga. 23

Returning to the size of the tonal inventory, the eight-tone view gains support even from authors who present the six-tone view in their work. For instance, in his standard book on Vietnamese phonology, Doan 1977 presents the six-tone view as a formal, official view. However, in his discussion, he says that he makes this choice because the six-tone view is very popular and familiar in the Vietnamese literature, and because it is reflected in the orthography. After presenting proposals on different tonal inventories including the six-tone and eight-tone hypotheses, he comments that the eight-tone inventory is probably the best, because it reflects the traditional classification in poetry before the orthography was created. While one must question this argument on synchronic grounds, this does point to there being some diachronic evidence for the eight tone hypothesis.

When we look at tone from a different perspective, namely, the domain of tone, the eight-tone view also receives some support. This issue will be discussed in Chapter Six.

4.2.4. Summary

In this section I have shown that the eight-tone hypothesis has advantages over the six-tone hypothesis. Under my assumptions about the phonology of tone, the eight-

---

23 Such changes happen in other languages. For instance, in Punjabi the loss of original voiced aspirates and [h] led to a tonal distinction, while the same feature was responsible for the breathy-voiced vowels of Punjabi (Anderson 1978:136).
tone view provides a simpler account of neutralization in reduplication. In favour of the eight-tone hypothesis, I have also shown support from the historical development of tones. Only Cao 1998 argues that the reduplication process supports the eight-tone hypothesis. Other comments are mainly based on the historical development of Vietnamese tones and patterns of tones in traditional poetry. While more synchronic evidence is needed to support the eight-tone view over the six-tone hypothesis, I continue to assume the eight-tone hypothesis.

4.3. Formal structure of the two tones in obstruent-final syllables

In this section I present the formal structures of sac2 and nang2 under the eight-tone hypothesis. I assume that these two tones are distinguished from all other tones by the presence of the feature [obstruent] in their phonological representations, and I give justification for the presence of this feature. Evidence for the organization and markedness relationships between sac2 and nang2 is also provided.

4.3.1. The feature [obstruent]

Recall that in stop-final syllables, only the two tones sac2 and nang2 can occur. Recall also that the manner of articulation of the final consonants is predictable from the tones. Ngang, huyen, hoi, nga, sac1 and nang1 predict that a final consonant will be a sonorant, sac2 and nang2 predict that a final consonant will be an obstruent. From the fact that these two tones occur only in stop-final syllables, and by adopting the eight-tone view that sac2 and nang2 predict final obstruents in final position, I assume this characteristic is encoded in the structure of sac2 and nang2 by a feature I call [obstruent]
for convenience. In this section, I argue that tones with the feature [obstruent] are more marked than those without this feature.

In languages, it is not unusual to find that obstruents are neutralized in the coda position. For instance, in Spanish and Italian, except s, only sonorants can occur in final position (Clements 1990a:312). This phenomenon supports the claim that obstruents are more marked than sonorants in this position. Support for this claim is also found in the Vietnamese final inventory and in clitics. We first look at the final inventory.

4.3.1.1. Final inventory

The phonetic final inventory from Chapter One is repeated in (59).

(59) Phonetic final inventory

```
   p  t  [c  kp  k]
   m  n  [ɲ  ɲm  ɲ]
   w  j
```

Under the eight-tone hypothesis, final stops and nasals are in complementary distribution and thus are allophones. In this section, I assume that final consonants are not specified for manner, and I suggest how they are represented underlyingly and on the surface. First, I assume the feature organization of segments as developed in Clements 1990b and Clements and Hume 1995, among others. In this model, under the root node, a consonant has a C-place node (60a), and a vowel, under the vocalic node, has a V-place node and an aperture node as its sister (60b). An aperture node specifies
vowel height. The difference between a vowel and a glide is that the former has the aperture node, which the latter lacks (60c). X is a place feature.

(60) Feature organization of segments

<table>
<thead>
<tr>
<th>a. consonant</th>
<th>b. vowel</th>
<th>c. glide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root</td>
<td>Root</td>
<td>Root</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-pl</td>
<td>vocalic</td>
<td>vocalic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>V-pl</td>
<td>Aperture</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If we assume that nasal and stop in final position are allophones, we can propose that there are two types of underlying segments in final position, one for consonants and another for glides. Based on (60), the representation of a final consonant is illustrated in (61) below with relevant features only.
(61) Representation of final consonants

<table>
<thead>
<tr>
<th>a. Underlying stop/nasal</th>
<th>Surface stop</th>
<th>Surface nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root</td>
<td>Root</td>
<td>Root</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-pl</td>
<td>C-pl</td>
<td>C-pl</td>
</tr>
<tr>
<td></td>
<td>[obst]</td>
<td>Nasal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b. Underlying glide</th>
<th>Surface glide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root</td>
<td>Root</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>vocalic</td>
<td>vocalic</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>V-pl</td>
<td>V-pl</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

In (61a) the surface final voiceless stops and nasals share the same underlying representation, a Root node with C-place as a dependent. On the surface, there are two types of consonants: one with the [obstruent] feature and one without. With the [obstruent] feature, the consonant is realized as a voiceless stop. It gets the feature [obstruent] from the tone. Because I adopt the eight-tone hypothesis which assumes that this feature is a tone feature, [obstruent] is shown in the representations of sac2 and nang2 only. I do not attempt to show how [obstruent] is inserted here. In a second type,
I assume that the feature [nasal] is inserted by a default rule (e.g., Rice 1993), yielding a surface nasal.

In (61b) I assume that glides [w] and [j] have the underlying V-place node under the vocalic node. I assume that, unlike in nasals, the Nasal Default rule does not apply in glides, resulting in a plain sonorant. This can be considered to be the result of a constraint that prohibits a node from being specified for nasal if the vocalic node is present in the segment (note that there are no nasalized vowels in Vietnamese).

My analysis leads to the conclusion that there is no underlying manner feature for the syllable-final obstruents or sonorants, the surface feature for sonorants is a default feature inserted at the phonetics. Thus, sac2 and nang2 force the insertion of obstruent.

4.3.1.2. Surface clitics and markedness

Clitics provide evidence that in Vietnamese, sonorants are unmarked in the syllable-final environment. In this section I show that in clitics the coda position is unspecified for manner, and [nasal] is a default feature. Some examples of clitics are given in (62). Only the forms involved in cliticization are shown, not the whole phrases. The clitics are underlined (see Chapter Two, section 2.3.6 for more discussion).

(62). Examples of clitics

a. [haːt] (sac2) + [laːm] (huyen) > [haːt] (sac2) + [n] (huyen)
b. [kɔ] (sac1) + [mot] (nang2) > [kɔ] (sac1) + [ɔ] (huyen)
(62) shows the process of cliticization. We now focus on the segments, including consonants, glides and vowels. Recall from Chapter Two, section 2.3.6, and as we can see in (62), that after cliticizing onto the host, all segments of the clitic disappear, only the tone remains. The clitic surfaces as a nasal if the host ends in an obstruent (62a). Note that in (62a) it just so happens that the clitic [lam] ends in a nasal. In this case, the nasal has the place of articulation of the final obstruent, i.e., it is [n] in (62a) since [ha:t] ends in [t]. The clitic’s tone is realized on the nasal. If the host ends in a vowel (62b) or a glide (62c), the vowel or glide is lengthened and the clitic’s tone is realized on the lengthened part. Note that the clitic [mot] in (62b) occurs with huyen, not the original nang2 tone. The clitic [mot] is discussed in Chapter Two, footnote 10. The main point is that, no matter what tone the base has, the clitic must have an unmarked register tone that predicts a sonorant.

The fact that only sonorants (nasal, vowel and glide) can surface in the clitics supports the claim that in that environment, sonorants are less marked than obstruents.

In summary, I have shown that the complementary distribution in the final inventory along with the surface form of clitics suggest that obstruents are more marked than sonorants in these environments. Recall that by adopting the eight-tone view, I assume that the surface final stop in the coda position is a tonal feature of sac2 and nang2. In the following section I present the structure of sac2 and nang2.
4.3.2. Representation of sac2 and nang2

To represent these tones structurally, I assume that the feature [obstruent] is part of these tones, and that the predictability of a final stop with these tones follows from its presence. This feature is a marked feature in this position, therefore, as assumed in Chapter Two, section 2.2, it is present underlyingly in the representation of sac2 and nang2.

(63) Phonological representations of sac2 and nang2

\[
\begin{align*}
\text{sac2} & \quad \text{nang2} \\
\text{C} & \quad \text{C} \\
[\text{obstruent}] & \quad [\text{obstruent}] \\
\text{•} & \quad \text{•} \\
\end{align*}
\]

In (63), sac2 and nang2 have the feature [obstruent] under the contour node, because these are the only tones which occur with stop-final syllables. Further work is needed to see whether [obstruent] is best under Contour or under the tonal node. Before I discuss the structure of sac2 and nang2, I repeat the structure of sac1 and nang1 from Chapter Two.
If we compare (63) and (64), the only difference under the contour node is the presence of [obstruent] in sac2 and nang2. Like sac1 and nang1, sac2 and nang2 have two points specified under the feature [obstruent] to show the movement. Like sac1, sac2 does not have a register feature because it is unmarked for register. On the register side, like nang1, nang2 is marked with the register feature [laryngeal]. However, unlike nang1, [laryngeal] in nang2 has [spread] as its dependent. Justification of the register feature is given in Chapter Five.

In the structure in (63), while sac2 and nang2 are equally complex on the contour side, sac2 is less complex than nang2 on the register side with the feature [spread]. The restricted distribution of sac2 and nang2 can be taken as an indicator of the markedness of these tones.

If we compare the structure of sac2 and nang2 in (63) with the structure of the other six tones from Chapter Two, section 2.2.2., repeated here as (65), we see that overall, sac2 and nang2 are the most complex tones vertically with the feature [obstruent] under the contour node.
(65) Formal structure of six tones in sonorant-final syllables

a. \(\text{ngang}\) \hspace{1cm} \(\text{huyen}\)

\[
\begin{array}{c}
\text{C} \\
\text{C} \\
\bullet \\
\bullet \quad \text{[laryngeal]} \\
\text{[spread]}
\end{array}
\]

b. \(\text{sacl}\) \hspace{1cm} \(\text{nang1}\)

\[
\begin{array}{c}
\text{C} \\
\text{C} \\
\bullet \\
\bullet \quad \text{[laryngeal]} \\
\text{[constricted]}
\end{array}
\]

c. \(\text{hoi}\) \hspace{1cm} \(\text{nga}\)

\[
\begin{array}{c}
\text{C} \\
\text{C} \\
\bullet \\
\bullet \quad \text{[laryngeal]} \\
\text{[constricted]}
\end{array}
\]
4.3.3. Evidence for the structure of sac2 and nang2

In this section I provide evidence for the register feature of sac2 and nang2, and show that nang2 is more complex than sac2.

4.3.3.1. Reduplication: evidence for the register feature of sac2 and nang2

In the process of reduplication discussed in Chapter Two, section 2.3.1, we saw that tones in the unmarked register pattern together in a group (ngang, sac1, hoi), and tones in the marked register form a group as well (huyen, nang1, nga). Recall that in this process, the register of the base is always replicated in the reduplicant. The data given earlier in (56), repeated in (66) with some more data, shows this process with sac2 and nang2. The base tone is underlined.

(66)

a. UNMARKED REGISTER

\[
\begin{align*}
\text{sac} & \quad (\text{sac2}) \quad '\text{sharp}' \quad > \quad \text{sang sac} & \quad (\text{ngang-sac2}) \quad '\text{rather sharp}' \\
\text{mat} & \quad (\text{sac2}) \quad '\text{cool}' \quad > \quad \text{man mat} & \quad (\text{ngang-sac2}) \quad '\text{rather cool}' \\
\text{xop} & \quad (\text{sac2}) \quad '\text{spongy}' \quad > \quad \text{xom xop} & \quad (\text{ngang-sac2}) \quad '\text{rather spongy}' \\
\end{align*}
\]

b. MARKED REGISTER

\[
\begin{align*}
\text{dep} & \quad (\text{nang2}) \quad '\text{beautiful}' \quad > \quad \text{dem dep} & \quad (\text{huyen-nang2}) \quad '\text{rather beautiful}' \\
\text{ngot} & \quad (\text{nang2}) \quad '\text{sweet}' \quad > \quad \text{ngon ngot} & \quad (\text{huyen-nang2}) \quad '\text{rather sweet}' \\
\text{sach} & \quad (\text{nang2}) \quad '\text{clean}' \quad > \quad \text{sanh sach} & \quad (\text{huyen-nang2}) \quad '\text{rather clean}' \\
\end{align*}
\]
In section 4.2.1 I stated that the final stop in the base changes to the homorganic nasal in the reduplicant. Ignoring that fact in (66a), where the base is sac2, we see that the reduplicant is always ngang, an unmarked register tone. In (66b), the base is nang2, the reduplicant is always huyen, a marked register tone. In this process, sac2 patterns as if it were in the same group as sac1 and hoi, i.e., they all have ngang in the reduplicant. Nang2 patterns as if it were in the same group as nang1 and nga, i.e., they all have huyen in the reduplicant. This fact shows that sac2 and nang2 share the same register feature with other unmarked and marked register tones, respectively: as proposed in (63), sac2 does not have the register feature; nang2 has the marked feature [laryngeal] with [spread] as its dependent, like in huyen. The use of the feature [spread] will be justified in Chapter Five.

In summary, the behaviour of sac2 and nang2 in reduplication shows that sac2 is an unmarked register tone and nang2 is a marked register tone. If we include sac2 and nang2 in the overall picture, the markessness of tonal registers is summarized as in (67).

(67) Markedness of tonal registers

<table>
<thead>
<tr>
<th>Unmarked register</th>
<th>ngang</th>
<th>sac1</th>
<th>hoi</th>
<th>sac2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marked register</td>
<td>huyen</td>
<td>nang1</td>
<td>nga</td>
<td>nang2</td>
</tr>
</tbody>
</table>

4.3.3.2. Borrowings: evidence for sac2 being less marked than nang2

Borrowings provide evidence that sac2 is less complex, and therefore less marked, than nang2. Borrowed forms which end in non-sonorant segments can only occur with a final stop. Consequently, the two possible tones in this position are sac2
and nang2. Some examples of borrowings are given in (68), taken from Nguyen 1975 and Nguyen 1998. The tone sac2 is represented by the diacritic ' ´ ' above the vowel, and nang2 is represented by the diacritic '.' underneath the vowel. Among the borrowed forms in (68), some can occur with either sac2 or nang2, i.e., (68h), (68i), (68l). The data in (68) shows sac2 is more common than nang2 in borrowings.

(68) Borrowings occur with sac2 or nang2

a. xép (Fr. chef) (sac2) ‘chief’

b. dîp (Fr. jupe) (sac2) ‘short skirt’

c. xà lách (Fr. salade) (huyen-sac2) ‘salad’

d. cà rốt (Fr. carotte) (huyen-sac2) ‘carrot’

e. gác (Fr. garde) (sac2) ‘guard’

f. séc (Fr. chèque) (sac2) ‘cheque’

g. mù tạc (Fr. moutarde) (huyen-nang2) ‘mustard’

h. gác dò bu (Fr. garde-boue) (sac2-huyen-ngang) ‘mudguard’
or gác dò bu (nang2-huyen-ngang)

i. xiéc (sac2) or xiếc (nang2) (Fr. cirque) ‘circus’

k. soác (sac2) or soôc (nang2) (Eng.) ‘short’

l. các vi dít (Fr. carte de visit) (sac2-ngang-sac2) ‘business card’
or các vi dít (nang2-ngang-sac2)
Interestingly, (68I) does not have nang2 in both syllables of the word, i.e., *cạc vi dît (nang2-ntag-nang2), and if only a single form is available, it almost has sac2, not nang2 (except for (68g), which is rare). For instance, lúp/* lúp (Fr. loupe), tip/*tip (Fr. type), xúp/*xúp (Fr. soupe), ráp/*rap (Eng. rap music), xęp/*xęp (Fr. chef), róc/*roc (Eng. rock music), a-xít/ *a-xít (Fr. acid), tuóc-no-vít / *tuóc-no-vít (Fr. tournevis), xà lách/*xà lách (Fr. salade), phât phú/ *phát phút (Eng. fast food).

The dominance of sac2 over nang2 in this type of borrowing has led researchers to claim that only sac2 occurs in this syllable type (Burton 1992). However, nang2 is also found, as we have seen. The more frequent distribution of sac2 over nang2 in this type can be explained if sac2 is less marked than nang2.

4.3.3.3. Frequency of occurrence: sac2 is less marked than nang2

Frequency effects also lend some support to the claim that sac2 is less marked than nang2, and to the eight-tone hypothesis. As we have seen from (28) in Chapter Two, section 2.3.3, repeated below as (69). For instance, the pair nang and huyen occur more frequently than the pair sac1 and nang1.

(69) Frequency of tones in Vo (1997) excluding -p, -t, -k

<table>
<thead>
<tr>
<th>Total</th>
<th>nang</th>
<th>huyen</th>
<th>sac1</th>
<th>nang1</th>
<th>boi</th>
<th>nga</th>
</tr>
</thead>
<tbody>
<tr>
<td>4243</td>
<td>1029</td>
<td>840</td>
<td>845</td>
<td>606</td>
<td>570</td>
<td>353</td>
</tr>
</tbody>
</table>
However, if \textit{sac} and \textit{nang} in both syllable types are regarded as the same, giving only six tones in the system, the frequencies are unexpected. In (70), Vo 1997 shows the actual numbers of tones in which \textit{sac} and \textit{nang} include both types of syllables: stop and non-stop finally.

((70) Frequency of tones including \textit{-p, -t, -k} (Vo 1997))

<table>
<thead>
<tr>
<th>Total</th>
<th>sac</th>
<th>nang</th>
<th>ngang</th>
<th>huyen</th>
<th>hoi</th>
<th>nga</th>
</tr>
</thead>
<tbody>
<tr>
<td>4243</td>
<td>1426</td>
<td>1045</td>
<td>1029</td>
<td>840</td>
<td>570</td>
<td>353</td>
</tr>
</tbody>
</table>

In (70), \textit{ngang} and \textit{huyen}, the two unmarked tones, as agreed in all analyses, have a lower frequency than \textit{sac} and \textit{nang}. Assuming that the less marked the tone, the more frequently it occurs, the markedness relationship between pairs would be odd: given the fact that \textit{ngang} and \textit{huyen} are usually claimed to be unmarked by their patterns in reduplication and in borrowings, we would expect them to occur more frequently than \textit{sac} and \textit{nang}. However, they occur less frequently than \textit{sac} and \textit{nang}. The six-tone hypothesis is unable to explain this fact.

From Tables (69) and (70), the frequency of \textit{sac2} and \textit{nang2} (in stop-final syllables) is deduced as in (71).

((71) Frequency of \textit{sac2} and \textit{nang2} in \textit{-p, -t, -k})

\textit{sac:} 581 \hspace{1cm} \textit{nang:} 439
With the numbers of occurrence of $sac2$ and $nang2$ in (71), $nang2$ occurs less frequently than $sac2$, confirming the claim that $nang2$ is more complex than $sac2$. Moreover, $sac2$ and $nang2$ occur less frequently than $ngang$ and $huyen$, the two unmarked tones in all analyses. Note that, however, the hypothesis which separates $sac$ and $nang$ in stop-final syllables from those in sonorant-final syllables does not compare the two tones in stop-final syllables with those in sonorant-final syllables, since they are in different environment.

I summarize the frequency of occurrence of eight tones in (72). The bold border separates the two tones in stop-final syllables from the rest.

\[
\begin{array}{llllllll}
\text{ngang} & \text{huyen} & \text{sac1} & \text{nang1} & \text{hoi} & \text{nga} & \text{sac2} & \text{nang2} \\
1029 & 840 & 845 & 606 & 570 & 353 & 581 & 439
\end{array}
\]

The table in (72) shows that among tones in sonorant-final syllables, the unmarked tones, $ngang$ and $huyen$, occur the most frequently (1029 and 840 respectively) with the ranking of markedness among tones shown in (73). The ranking in (73) summarizes the ranking of six tones in sonorant-final syllables in (21) and (22) in Chapter Two, adding the two tones in stop-final syllables. ‘$<$’ means ‘less marked than’.

\[
\text{ngang} < \text{huyen}, \text{sac1} < \text{nang1}, \text{hoi} < \text{nga}, \text{sac2} < \text{nang2}
\]
If frequency of occurrence can be a diagnostic for the markedness of tones, the ranking of markedness of tones in (73) predicts that among tones in sonorant-final syllables, the least marked tone, ngang, has the highest frequency and the most marked one, nga, has the lowest frequency. Among pairs, the least marked pair ngang-huyen occurs the most and the most marked one hoi-nga occurs the least. Between the two tones in stop-final syllables, nang2 occurs less frequently than sac2. These predictions about frequency in Vietnamese are borne out, providing evidence for the markedness relations proposed here.

The rarity of sac2 and nang2 in the six-tone view is interpreted as the rarity of final -p, -t, -k. However, if we assume there are eight tones in the system, it is due to the rarity of the two tones sac2 and nang2, because they are regarded as tonal features.

4.4. Summary

In this chapter I presented the two major views concerning the tonal inventory of Vietnamese. I show that although the six-tone view is the standard hypothesis, the eight-tone hypothesis has certain advantages over the six-tone hypothesis. I also provided more acoustic evidence to support the eight-tone hypothesis. Reduplication, borrowings and frequency of occurrence are shown compatible with the claim about the markedness relationship between the two tones sac2 and nang2, i.e., sac2 is an unmarked register tone and nang2 is a marked register tone, or sac2 is less marked than nang2.
Chapter 5

THE PHONETICS-PHONOLOGY INTERFACE

5.0. Introduction

In this chapter, I discuss the implications of the acoustic results presented in Chapter Three, and the relationship between the phonetics and phonology of Vietnamese tones.

In particular, in section 5.1 I discuss the implications of the experimental study for the phonology of tone. I argue that the mismatch problem between hoi and nga discussed in Chapter Two is an illusory one. I also present evidence for the laryngeal features proposed in Chapter Two. The study shows that in Vietnamese, among various phonetic features, tonal shape and the phonation types of creakiness and breathiness are the most reliable, and are primary features for differentiating the tones. I argue that pitch height can be predicted from tonal shape and phonation types. In section 5.2 I discuss the mapping between the phonetics and phonology of tones, how phonetic features are represented, and how phonological features are realized. Variation within individuals and across speakers is also discussed.

The characteristic features of each tone are found in Table 4. The parentheses around the feature curve in hoi show that this feature varies from speaker to speaker, as discussed in Chapter Three. The last row (Pitch of each tone) does not show 'pitch' in the
way it is traditionally classified. Rather it shows how pitch height is predicted from Register, e.g., breathiness predicts lowness in the tone, either throughout as in *huyen* or at certain point as in *hoi*. This point will be made clear when we discuss phonation types in section 5.1.2.

Table 4: Phonetic characteristic features of tones

<table>
<thead>
<tr>
<th>Tone</th>
<th>nga</th>
<th>huyen</th>
<th>sac1</th>
<th>nang1</th>
<th>hoi</th>
<th>nga</th>
<th>sac2</th>
<th>nang2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contour</td>
<td>level</td>
<td>level</td>
<td>rise</td>
<td>fall</td>
<td>(curve)</td>
<td>curve</td>
<td>rise</td>
<td>fall</td>
</tr>
<tr>
<td>Register</td>
<td>modal</td>
<td>breathy</td>
<td>modal</td>
<td>creaky</td>
<td>breathy</td>
<td>creaky</td>
<td>modal</td>
<td>breathy</td>
</tr>
<tr>
<td>(Pitch)</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
</tbody>
</table>

5.1. Implications of the acoustic study for the phonology

Recall that *hoi*, traditionally classified as a phonetic low tone, patterns phonologically as a high tone; *nga*, traditionally classified as a phonetic high tone, patterns phonologically as a low tone (Chapter Two, section 2.3.1). The phonetic characteristics and phonological patterning of *hoi* and *nga* (in shaded cells) from previous chapters are repeated below in tables 5 and 6.

Table 5: Phonetic classification of tones

<table>
<thead>
<tr>
<th>High</th>
<th>nga</th>
<th>sac1</th>
<th>nga</th>
<th>sac2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>huyen</td>
<td>nang1</td>
<td>hoi</td>
<td>nang2</td>
</tr>
</tbody>
</table>
Table 6: Phonological patterning of tones in reduplication

<table>
<thead>
<tr>
<th>Unmarked</th>
<th>ngang</th>
<th>sac1</th>
<th>hoi</th>
<th>sac2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marked</td>
<td>huyen</td>
<td>nang1</td>
<td>nga</td>
<td>nang2</td>
</tr>
</tbody>
</table>

In the following section I will show that hoi patterns phonetically as well as phonologically (as we have seen in Chapter Two) with the other high tones; nga patterns phonetically as well as phonologically (see Chapter Two) with other low tones.

5.1.1. Hoi is low and nga is high: the mismatch is illusory

The acoustic study shows that hoi is higher than nga. In some cases hoi is higher than nga only in the middle, the most crucial point in these tones. This is consistent with the pattern found in the previous studies discussed in Chapter Three, section 3.1. In other cases, hoi is higher than nga throughout. In this section I show evidence for each of these two patterns.

5.1.1.1. Hoi is partially higher than nga

I first present forms that have similar patterns to those in previous studies. In this pattern, there is a glottal stop in the middle of nga. If the glottal stop is not shown, hoi is higher than nga in the middle of the tone.

The glottal stop in the middle of nga is seen in the pitch tracks in Vu 1982 and in Nguyen and Edmondson 1997 discussed in Chapter Three, repeated here as Figures 29 and 30. Recall that the glottal stop is represented by a dotted line in Figure 29, and by a blank portion in Figure 30.
Figure 29. *Nga* has a glottal stop in Vu 1982

Normalized duration in decimal scale
Figure 30. Nga has a glottal stop in NE 1997

This pattern is also seen in my study. Figures 31 and 32 show a glottal stop in the middle of nga from Speakers Binh and Hung, respectively. In these figures, nga is broken by glottal stop, represented by a discontinuous portion in the middle of nga.
Figure 31. *Nga* has a glottal stop in [ta], [taj] from Speaker Binh
Figure 32. Nga has a glottal stop in [ma] and [ka] from Speaker Hung.
When the glottal stop is not present, *hoi* cuts across *nga* as seen in Hoang 1986, Chapter Three, and repeated here as Figure 33. This pattern is seen in speakers 3 and 4.

Figure 33. *Hoi* cuts across *nga* in Hoang 1986

This pattern is also seen in my study. Two forms are shown in Figure 34 from two female speakers, Binh and An, and in Figure 35 from two male speakers, Hoang and Son.
Figure 34. Hoi cuts across nga from Speakers Binh and An.
Figure 35. *Hoi* cuts across *nga* from Speakers Hoang and Son

**Hoang**. *hoinga-za*

![Graph showing frequency changes over time for Hoang.]  

**Son**. *hoinga-ka*

![Graph showing frequency changes over time for Son.]
In these cases, although both tones *hoi* and *nga* go down in the middle, the lowest part of *nga* is always lower than that in *hoi*. Previous studies have failed to notice this and have treated *nga* 'overall' as a high tone and *hoi* as a low tone ('overall' in the traditional usage is the average of an equally-spaced sample). I will show below that this view is simply inadequate.

There are two points that need to be clarified concerning F0 of tones. First, when one compares the F0 of *hoi* and *nga* - if comparing the height of F0 is a relevant thing to do – one should consider the relationship only of these two tones, e.g., *hoi* is lower or higher than *nga*. Outside this relationship, the definition of 'low-high' for these tones is not relevant. For instance, if tones are divided into their registers according to overall height in the pitch span (Cu et al 1977, Huu and Vuong 1980 and others, see definitions of 'register' in Chapter Two), then, both *hoi* and *nga* can be clearly specified as low tones (see Figure 36). In Figure 36 both *hoi* and *nga*, together with *nang1*, fall completely into the lower part of the pitch span. In this figure, I only show six tones, without *sac2* and *nang2*, to make it easier for the reader.
Figure 36. *Hoi* and *nga* are both in the low register in Speaker Dung.

The same thing can be said for the two forms [ta] and [ka] from Speaker Binh in Figure 37. Again, I only show six tones. *Sac2* and *nang2* are omitted.
Figure 37. *Hoi* and *ngä* in the low register in [ta], [ka] from Speaker Binh
In these forms, *hoi* and *nga* both fall into the lower part of the pitch span. This suggests that tones should be compared within their pairs, not in general with other tones. This is seen not only in my study but also in previous studies. For instance, if we look at Speaker 6 in Figure 30 from Nguyen and Edmondson 1997, the three tones that occur in the lower part are *ngang, huyen, hoi*, and the three that occur in the higher part (if we ignore the glottal stop in *nga*) are *sac1, nang1* and *nga*. Note that traditionally *ngang* is a high tone and *nang1* is a low tone, but *nang1* shares the same pitch space with high tones, and *sac1* with low tones. If a tone is classified as low or high according to which part of the pitch range it falls into, *nang1* should be classified as a high tone, and *ngang* a low tone in Figure 30.

The second point is that, depending on which point one looks at, *hoi* can be regarded as low and *nga* as high, or vice versa. Recall that *hoi* is breathy and *nga* is creaky. If we look at the middle part of *hoi* and *nga*, if *nga* is creaky in the middle, this part is lower in *nga* than in *hoi*, since creakiness makes the tone lower than breathiness does (Marasek 1997, Ladefoged and Maddieson 1996 among others). If *nga* has a glottal stop in the middle, then the F0 comparison is impossible, because there is no phonation during a glottal stop.

*Nga* is usually treated as a high tone because it has a very high end. This end makes the tone ‘overall’ high. This argument for treating *nga* as high pitch phonetically is not entirely convincing. First, as seen in Chapter Three, section 3.3.1.2, the end of *nga* can be high or low, varying from speaker to speaker, form to form, and token to token. Some forms are repeated in Figures 38 and 39. In Figure 38, the end of *nga* in the same form [ta] is low for Speaker Phuong but high for Speaker An.
Figure 38. The end of *nga* is low from Phuong, high from An in the same form [ta]
Figure 39. The end of nga varies in different tokens of [ma] from the same speaker
In Figure 39, in the same form [ma], nga has a low end in token II but a high one in token I from Speaker Binh. Therefore, a high ending of nga in the pitch range is not a good criterion to classify nga as a high tone.

While the highest of the endpoint of nga can vary, nevertheless something is constant: namely, contour (tonal shape) is a distinctive feature in Vietnamese tones. For instance, hoi and nga are the two curved tones. Nga without the middle part would be like sac1, a rising tone. Therefore, if one must compare hoi and nga, the middle part cannot be ignored. In the following section, we will see that in many cases, nga does not need to be high, as long as it is curved.

Nga is traditionally classified as a high tone because of the high pitch at the end. If tonal height is instead based on the 'overall shape', we can challenge the traditional classification of tones. For instance, sac1 is classified as a high-rising tone and huyen as a low-level tone in all accounts. However, if we look at Figure 40, there are some problems. Figure 40 shows pitch trajectories of these two tones, huyen and sac1, in the two forms [ka] and [sa]. Huyen starts higher than sac1 and gradually goes down. Sac1 starts lower than huyen, goes flat and only rises to be higher than huyen during the last 30ms. So if overall pitch height is a criterion to classify tonal height, the high-rising sac1 would be low and the low-level huyen would be high.
This pattern is seen in many other forms in the same speaker. We see that only relative tonal height within a contour pair is meaningful.
In summary, I have shown in this section that if the tonal heights of *hoi* and *nga* are compared, *nga* is always lower than *hoi* at the middle part of the tone. I turn now to cases where *hoi* is higher than *nga* throughout.

5.1.1.2. *Hoi* is higher than *nga* throughout.

So far we have seen the following relationship between *hoi* and *nga*: *hoi* is higher than *nga* in the middle. In another pattern, the whole of tone *hoi* is higher than *nga*. This pattern is found not just in one speaker but in various speakers, and even in different tokens from the same speakers. Figures 41, 42 and 43 show some data from speakers Dung, Phuong and Khanh, respectively. In every token, *hoi* is higher than *nga* throughout. Figures 41a, b show different tokens of the form [ta] (Figure 41a) and [ka] (Figure 41b) from Speaker Dung. The onset of *nga* in some cases is higher than that of *hoi*, but the onset is not crucial, as we have seen from Chapter Three, section 3.3.3.3.
Figures 41a. *Hoi* is higher than *nga* throughout in [ta] in different tokens from Dung.
Figures 41b. *Hoi* is higher than *nga* throughout in [ka] in different tokens from Dung.
Speaker Phuong also has *hoi* higher than *nga* throughout. Figures 42 shows [ka] (42a) and [ta] (42b), from Speaker Phuong. Here we see the same pattern in another speaker: *hoi* is higher than *nga* throughout.

Figure 42a. *Hoi* is higher than *nga* throughout in [ka] from Speaker Phuong.

![Graph showing the comparison of *hoi* and *nga* over time in Speaker Phuong's speech.](image)
Figure 42b. *Hoi* is higher than *nga* throughout in [ta] from Speaker Phuong

Figure 43 shows data from Speaker Binh, again for the same form, [ta], with different tokens. Although there is a glottal stop in the middle of *nga*, the whole tone *nga* is lower than the tone *hoi*. 
5.1.1.3. Summary

In this section I have shown that *hoi* is clearly higher than *nga* in its fundamental frequency in many cases. This result challenges the view that *hoi* is low and *nga* is high. If tonal height is defined according to the maximal low/high points, the lowest point of *nga* is always lower than that of *hoi*. The middle point is important in these two tones, because without a curve, *nga* is similar to *sac1*, a high-rising tone, and *hoi* would be confused with *huyen*, a low-level tone.

These findings are important in challenging the claim that there is a mismatch between the phonetics and phonology of *hoi* and *nga*. The phonetic basis that is assumed
in the *hoi*/*nga* pair - *hoi* is low and *nga* is high - is simply not a true claim about the phonetics of these tones. One might think we could stop here, and simply claim that the phonological diagram in Table 3 accurately represents the phonetics as well. However, this is not actually so.

I argued in Chapter Three, section 3.3.3.1, that F0 is not a reliable feature to distinguish register pairs because it varies from speaker to speaker and from token to token. In addition, in many cases there is no pitch height difference between register pairs. Instead, tonal shape and phonation types, which will be discussed in the next section, emerge as the primary cues to distinguish register pairs in Vietnamese tones. For instance, I will show in section 5.1.2.2 that in the two curved tones *hoi* and *nga*, the middle point is crucial and invariant, but the end point can be varied.

In summary, I have provided evidence that even in a theory employing tonal height, *hoi* can be higher than *nga*. Therefore, the mismatch between phonological height and phonetic height claimed in the literature is not found when we carefully examine various aspects of the phonetics of tone.

5.1.2. Phonation types: a motivation for register features

One of the criteria to determine whether a tonal feature is phonological is that the feature should be able to group tones into natural classes, and thus explain tonal patterning found in the language. Assuming features are actual articulatory instructions (Rice and Avery 1991) along with the theory of Contrastive Specification (see Chapter One, section 1.3), we would expect a phonological feature to be reliable or a constant part of the phonetic realization of a segment. A distinctive feature is not something that
optionally occurs or freely varies. In other words, the phonological feature should play a
role in producing a contrast in the language, and should have a regular mapping into
the phonetics.

We have seen from Chapter Two, section 2.4.1, that none of the phonetic features
proposed in the Vietnamese literature properly groups tones into natural classes. For
instance, glottalization (i.e., glottal stop or creakiness) would group sac1 and nga into a
pair, but these two tones do not pattern together. Tonal height seems to be the most
reasonable feature in grouping tones into two registers according to their pattern in
reduplication. However, even with the mismatch problem between phonetics and
phonology solved, F0 of tone is unstable, leading me to question its status as the
appropriate phonological feature to classify tones into natural classes in Vietnamese.

In this section, I provide justification for the register feature [laryngeal] proposed
in Chapter Two. I show that the phonation types of breathiness and creakiness are stable
for all speakers. More importantly, these features group tones into classes that allow a
natural statement of phonological rules (Yip 1995). I also argue that in Vietnamese tonal
height is derived from tonal shape and phonation types.

First I repeat below Table 4, page 175, with contour and register features.

<table>
<thead>
<tr>
<th>Tone</th>
<th>ngang</th>
<th>huyen</th>
<th>sac1</th>
<th>nang1</th>
<th>hoi</th>
<th>nga</th>
<th>sac2</th>
<th>nang2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contour</td>
<td>level</td>
<td>level</td>
<td>rise</td>
<td>fall</td>
<td>(curve)</td>
<td>curve</td>
<td>rise</td>
<td>fall</td>
</tr>
<tr>
<td>Register</td>
<td>modal</td>
<td>breathy</td>
<td>modal</td>
<td>creaky</td>
<td>breathy</td>
<td>creaky</td>
<td>modal</td>
<td>breathy</td>
</tr>
<tr>
<td>(Pitch)</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
</tbody>
</table>

199
In particular, I will show that although they differ in degree, huyen, hoi and nang2 are always breathy, and nang1 and nga are always creaky. I will also show that when there is breathiness or creakiness at a certain point in a contour, the pitch height is low at that point. If breathiness or creakiness is spread throughout a tone span, the pitch height is low throughout. However, note that, as we have seen in Chapter Three, the onset of tone is not critical in Vietnamese. In this language the distinctive part of a tone does not start until after approximately the first 60ms at the earliest.

5.1.2.1. Phonation types are stable features

In the previous chapter and in section 5.1.1, we saw that F0 of tone varies from token to token, and from speaker to speaker. For instance, nga can have a very high endpoint for one speaker but a low one for another: the endpoint can even differ from token to token for the same speaker. However, when we look at phonation type, it is quite consistent both within individuals and across speakers. As we saw in Table 4, tones that have non-modal phonation are huyen, nang1, nga, nang2, and hoi. Huyen, hoi and nang2 are breathy. Nang1 and nga are creaky. In this section I focus only on non-modal tones, i.e., tones that have breathiness or creakiness. I show that creakiness and breathiness are always found in the marked tones, i.e., lower register tones. Among tones with phonation types, note that hoi is unmarked for register but it is breathy. Breathiness in hoi will also be discussed.

First let us look at huyen. Breathiness in huyen is the only difference between the two level tones, ngang and huyen. For all speakers, huyen always has breathiness. However, the degree of breathiness in huyen varies from speaker to speaker. Figure 44
shows the waveforms and spectrograms of *ngang* and *huyen* from a male, Speaker Hung. Here the tone *huyen* is breathy, but only minimally. The spectrogram of *huyen* is shown in the last row. Although the glottal pulses are periodic in the waveform, breathiness is seen as reduced amplitude after 70 ms, increasing after 150 ms. Also, due to the incomplete closure of the vocal folds, we can see some space between glottal pulses in the spectrogram, especially after 150 ms. However, these signals are not very strong.

Figure 44. Breathiness in *huyen* in the form [ma] from Speaker Hung

Figure 45 shows the spectrograms of *ngang* and *huyen* of the same form [ma] from another male, Speaker Son. For this speaker, *huyen* is very breathy.
In Figure 45, the waveform and spectrogram of huyen are in the second and last rows, respectively. Breathiness in huyen is clearly seen in the waveform, represented by very weak cyclical pulses, especially after 150 ms. In the spectrogram, there are gaps between pulses towards the end of the vowel. Interestingly, if we compare the spectrograms of the two tones ngang and huyen in Figures 44 and 45 with their corresponding pitch graphs, the difference in F0 between ngang and huyen in Speaker Son is much smaller than that in Speaker Hung. The corresponding pitch graphs of ngang and huyen from Speakers Hung and Son are shown in Figures 46 and 47, respectively.
Figure 46. Pitch graph of ngang and huyen in [ma] from Speaker Hung
This suggests that the speaker might have a choice: higher F0 for ngang, or stronger breathiness for huyen to the degree necessary to make a phonetic contrast. This suggests a trade-off between the two phonetic realizations, which in turn suggests some interesting speech perception experiments as a topic for future study. Whatever the choice, huyen always has some degree of breathiness. Recall that ngang does not have breathiness; it is modal voice. As expected, huyen with breathiness is lower than ngang.

Nang2 is another tone with breathiness. Recall that sac2 and nang2 occur only with final obstruents, which make the tone very short, as we saw for Speaker Son in Chapter Three. Due to shortness, this tone does not provide the length necessary for breathiness to be fully realized, but breathiness still can be identified. Figure 48 gives the spectrogram of sac2 and nang2 in the form [ta:k] from An, a female speaker.
As mentioned in Chapter Three, phonation types usually begin after approximately 60ms. In Figure 48 we see this in nang2 in the last row. The cycles are regular and periodic in sac2 in the third row since the tone has modal voice, while the cycles of nang2 become very weak with noise in high frequencies towards the end.

Notice that unlike nang1, there is no glottal stop in nang2. The fact that it is breathy and it does not have a glottal stop makes this tone look much like huyen, a level breathy tone. This is very noticeable when we look at the pitch graphs of these two tones, shown in Figure 16b of Chapter Three, and repeated here as Figure 49. The two lines are identical in height.
Figure 49. Huyen and nang2 in [ta] and [ta:k] from Speaker Dung

In other pitch graphs, these two tones are also very close together. This fact provides extra evidence that nang2 is as breathy as huyen. The only phonetic difference between nang2 and huyen is that in nang2 the vocal folds suddenly stop vibrating; there is oral closure, and the tone ends abruptly.

With the two other marked tones nang1 and nga, there is glottalization (either a glottal stop or creakiness) in every token as discussed in Chapter Three and in the following section.

In summary, while they can vary in degree (breathiness in huyen and nang2), specific phonation types are always seen in marked tones.
5.1.2.2. Tonal shape and phonation types are distinctive and account for tonal height

In this section, I argue that in Vietnamese phonation type functions distinctively and that pitch register is derived from tonal shape and phonation type. Particularly, I will show that creakiness and breathiness are laryngeal features which differentiate tones. A curved-shape tone predicts that the tone will go down in the middle (recall that Vietnamese curved tones are concave, i.e., tone falls and rises). Breathiness and creakiness predict lowness. Non-curved tones with modal voice are non-low.

Assuming that the feature that is invariant is distinctive, let us first look at nang1 and nga, two tones with laryngealization. I will argue that creakiness in these tones is distinctive. Height and length are not important as long as creakiness is present in the tone.

Contrary to the claim in the literature that nang1 ends with a glottal stop (see Chapter Three, section 3.2.2.2), there are two possibilities: the glottal stop can end the tone, or the vocal folds might continue to vibrate after that. If the vocal folds stop vibrating after glottal stop, the tone is very short. If the vocal folds still vibrate after the glottal stop, as a result, there is a portion of creakiness. This variation after the glottal stop in nang1 shows that the portion after the glottal stop is not distinctive in this tone. Figure 50 shows the pitch graph of nang1 in the form [ka] from speakers Binh and An. Recall that nang1 is written as na in the legend of the pitch graph.
In Figure 50, *nang1* is short in Speaker An, ending at around 120 ms. However, in Speaker Binh the tone goes further and ends at approximately 210 ms. The difference in length is shown in the corresponding spectrograms in Figure 51. In this figure, the spectrogram of *nang1* in Binh is in the last row. In the third row (An), the vocal folds stop vibrating after 120 ms. In Binh in the last row, creakiness starts after 120 ms. The distances between pulses are increased to the longest, then decreased after 180 ms toward the end of the vowel.
Interestingly, we can find a similar situation in the same speakers with different forms, and even in the same form but different tokens within the same speaker. Figure 24 shows the pitch graph of \textit{nang1} in the form [ma] with two different tokens from Speaker An. In this graph \textit{nang1} in token I, represented by a line with grey diamonds, stops at about 120ms. The tone has about the same F0 in token III, except that after the glottal stop, it goes farther ending at around 200ms.
Figure 52. *Nang1* in [ma] can be short or long in the same speaker.

Again, the difference in length of *nang1* in Figure 52 is shown in their corresponding spectrograms, given in Figure 53.
Figure 53. Spectrograms of *nang*1 in [ma] from Speaker An with different tokens

In Figure 53, the spectrogram of the first token of [ma] is in the third row. *Nang*1 is very short. After the glottal stop, the vocal folds stop vibrating for 60 ms. Before vibration fully stops the vocal folds vibrate very weakly, but do not produce a clear cycle. In the spectrogram of the third token in the last row, after 120 ms the vocal folds vibrate intermittently and produce increased distances between pulses, followed by a few decreased distances of pulses toward the end.

Assume that phonation predicts tonal height. The glottal stop pulls the tone down suddenly, as seen in Figures 52 and 53, and the creaky portion makes the tone fall
and become low, as in the long [ka] from Binh in Figure 51. The prediction that creaky
tones will be low explains why nang1 in [ma] Figure 52 is very high during the first
90ms before the glottal stop occurs: this part is high, about 270 Hz, because it has modal
voice. This part is even higher than ngang and sac1, traditionally high level and rising
tones, respectively. This can be is seen in Figure 54, showing six tones from Speaker An.
Sac2 and nang2 are not important for our focus here and are omitted for simplicity. In
this figure, if we ignore the last 30ms in nang1 where the tone is pulled down by the
glottal stop, the level part of nang1 is higher than the whole tones ngang and sac1. This
creates a paradox: nang1, a traditional ‘low’ tone in every description, is higher than two
unquestionably ‘high’ tones, ngang and sac1.
Without referring to the glottal stop, a theory that employs pitch height as a distinctive feature cannot explain why nang1, a low tone traditionally, can be higher than high tones. However, if we assume that phonation type is the main cue, then as long as the glottal stop/creakiness is present, the tone is perceptible regardless of the height of the portion before the glottal stop. This explains why in some cases nang1 can stop after the glottal stop (this is a common description of nang1 in traditional literature) or continue after that with creakiness in other cases.

This phenomenon is seen not only in my study but also in Nguyen and Edmondson's results. Figure 55 repeats Figure 30, showing the pitch graphs of six tones.
in open syllables from Speaker 6. In this figure, *nangl* starts as high as *nga*. It goes down a little and ends at about 120 ms. The total tone, therefore, is considerably higher than *ngang*, a high level tone.

Figure 55. *Nangl* is higher than *ngang* in NE 1997

The shortness in *nangl* compared to other tones suggests that the glottal stop comes early in the tone and the vocal folds stop vibrating afterwards. Perhaps there is absolutely no voicing after the glottal stop and therefore, nothing that can be measured. The segmental portion without voicing cannot be seen in a pitch graph.

Again, a theory that treats pitch height as the primary feature of tone cannot explain this fact. However, if the laryngeal feature of creakiness is indeed a primary feature of tones, it can offer an account for this phenomenon. What is important is that there is a glottal stop (or creakiness) in this tone. As long as a tone has a glottal feature,
the tone is recognizable, no matter how high it is. However, the glotal stop will tend to lower the height of pitch.

Interestingly, the problem of the inconsistent description of nang1 is discussed in early acoustic studies of Vietnamese tone. Perhaps one of the earliest experimental works was conducted by Andreev and Gordina 1957. Vuong and Hoang 1994 point out that there is an inconsistency between the pitch height and shape of tone nang1 in Andreev and Gordina 1957 and in Haudricourt 1972 (cited in Vuong and Hoang 1994). Recall that Figure 29 shows the pitch graph of eight tones in Northern dialects (Vu 1982). The description of pitch height and shape of tones in this study are similar to that of Andreev and Gordina 1957 (cited in Hoang Tue and Hoang Minh 1975): in Vu 1982 and Andreev and Gordina 1957, nang1 is a low, falling tone. However, according to Haudricourt 1972, if we use numbers to represent tones, tone nang1 is 22, a low level tone, whereas huyen, the low level tone in all accounts, is 21. This fact suggests that, similar to the pattern of pitch height of nang1 shown in Figure 55 (NE), Haudricourt’s description captures only the first part of nang1 before the glottal stop. Without the glottal stop or the creaky portion, nang1 can have high pitch, as we have seen. This point supports the claim that in nang1, glottalization, not pitch height, is crucial. Glottalization cannot be represented in a pitch graph but only in a spectrogram, and this causes inconsistencies in the description of this tone in the literature. This point also shows that tonal shape (contour) has been very important in the Vietnamese traditional literature, i.e., whether nang1 is even (no glottal stop) or falling (with glottal stop).

I now turn to the curved tone nga, which is usually said to have a glottal stop in the middle. Recall that if there is a glottal stop, it breaks the tone into two pieces, as seen
in Figures 29 and 30. A similar pattern is also found in this study, as seen, for example, in Figures 31 and 32, where the glottal stop is indicated by a blank. In this case, there is no phonation at all as the vocal folds cease vibrating. There is no F0 to be measured at that point. If instead of a glottal stop, there is a creaky portion in the middle of the tone, the line representing this tone in the pitch graph is not broken but continuous. As we would predict, the creakiness makes the tone very low at that point, and the tone occurs with a clearly curved shape. After creakiness, the vocal folds either vibrate at very high frequencies, producing a clear V-shape tone as in Figures 34 and 35, or the vibration is not in very high frequencies, creating a slight curve as in Figure 41a or 42b.

Thus, nga can be very curved or slightly curved, but crucially, in each case there must be some creakiness in the middle of this tone. Let us look at this tone in Figure 56, from a male speaker.
Figure 56. Nga with a little creakiness from Speaker Hung

The first and second rows represent the waveforms of hoi and nga, respectively. Their corresponding spectrograms are the third and fourth rows, respectively. We focus on nga. Here we see that the tone nga does not have a glottal stop. It is also not strongly creaky. We do not find irregular glottal pulses or widely spaced pulses in the spectrogram. However, if we look closely at the waveform from approximately 50 ms to 90 ms, we see that the pulses have reduced intensity. The corresponding portion of the spectrogram shows some slight weakening of energy around 3200 to 4100 Hz. At low frequencies we see some distance between pulses from around 50 to 90 ms. This is
where creakiness occurs. In the pitch graph, we should see that the tone is low at this point since creakiness predicts lowness. Figure 57 shows the pitch graph of the spectrogram in Figure 56.

Figure 57. Pitch graph of tone nga in Figure 56 from Speaker Hung

![Pitch graph of tone nga](image)

When creakiness is not strong enough to make the tone go down in the middle, the tone is not clearly curved as a consequence (recall that curved tones are concave, not convex). In Figure 57, nga only goes down a little bit at approximately 60 ms and then rises. Although this type of realization of the tone nga is not very common, Figure 58 shows that it is not important whether nga is low or high, or how low/high the tone is. The crucial information is whether there is some creakiness in the tone. Without a glottal
stop or creakiness, nga would be confused with sac1, a rising tone. Figure 58 shows the pitch graphs of two tones, nga and sac1, from Speaker Hung. The tones are very similar in shape: both go down very slightly and then gradually rise.

Figure 58. Pitch graph of nga and sac1 from Speaker Hung

![Pitch Graph of nga and sac1](image)

Figure 59 gives the corresponding spectrograms. Interestingly, although nga and sac1 are similar in shape as seen in Figure 58, they are not exactly alike in their spectrograms. In Figure 59, the waveform of nga in the second row is of low intensity from 50 ms to 80 ms and we can see the corresponding portion in the spectrogram. Sac1, on the other hand, does not change in intensity at all throughout the tone. The waveform and spectrogram are quite stable throughout.
If we compare the waveform and spectrogram of the two tones in Figure 59 between 50 and 90 ms with the corresponding portion in Figure 58, we see that nga with a little creakiness is higher in F0 than sac1. This fact suggests that although creakiness is produced at low frequencies, low frequencies do not necessarily produce creakiness. This is very strong evidence that phonation predicts pitch height, not vice versa.

Figure 60 gives the complete pitch graph of eight tones in [sa]/[sa:k] from Hoang to show where sac1 and nga are in the pitch range of this speaker. Note that this speaker has the widest pitch range of the three male speakers. The value of F0 on the y-axis is
usually from 200 to 250 Hz, compared to 180 Hz in other males.

Figure 60. Eight tones in [sa] and [sa:k] from Speaker Hoang

From Figure 60, we see that *nga* does not fall as expected, but other tones are distributed as expected, e.g., *nang* falls, *hoi* falls, *sacl* rises and is higher than *ngang*. Note that this form gives another example where F0 is not very reliable: it is difficult to separate the two tones *huyen* and *nang2* in Figure 60. If we assume phonation types are distinctive and that breathiness lowers the tone, this fact can be explained. Because both *huyen* and *nang2* are breathy, they are very similar in F0. The difference is that *nang2* occurs in stop-final syllables, which shorten the tone. Phonologically this characteristic of *nang2* is captured by the feature [obstruent] (see Chapter Two and section 5.2 below).

In summary, without creakiness in the middle, *nga* would lose its most important
cue. Pitch height is not critical; we see that nga can have either a very high ending, making it partially higher than hoi (Figures 38, 39), or a very low ending, making it lower than hoi throughout (Figures 41, 42). When we treat the rising part of nga as phonetic, whether nga is 'lower' or 'higher' than hoi is no longer a problem, because it varies randomly. All that is required is that the F0 rises a little to make a curve.

So far, I have discussed the following tones with non-modal phonation types: huyen, nang1, nga and nang2. I have shown that in nang1 and nga, glottal stop or creakiness is consistently, and thus distinctively, present. In huyen and nang2 there is breathiness in the tone. Nang2 is distinguished from huyen by the feature [obstruent] that shortens the tone. All belong to the marked register. In the traditional literature, these tones are classified as low, except for nga, whose glottal stop is ignored in order to allow it to be grouped with high tones (see Table 2, section 5.1). However, in my account, phonation types are distinctive, therefore, glottal stop or creakiness in nga has to be taken into account. Assuming that creakiness and breathiness lower the tone, it is not surprising to see that all marked tones are often low, because each tone has either breathiness or creakiness at some point or throughout the tone. However, lowness, I argue, is derived from the phonation types of creakiness or breathiness, e.g., without creakiness or glottal stop, nang1, the traditionally low tone, appears as a high tone in terms of pitch range (Figures 54, 55). Pitch height, therefore, is not distinctive in my analysis. The fact that we can always find breathiness and creakiness in marked tones is phonetic motivation for the phonological feature [laryngeal] for these tones. This feature groups tones into a natural class, i.e., the marked register.

I now turn to hoi, the only tone that is breathy but unmarked for register. I will
argue that the breathiness in *hoi* comes from another requirement from the Contour side, and it is not a distinctive part of the tone.

The tone *hoi* is usually described as a curved tone that goes down and then up again, higher than the onset (see in this chapter Figure 29 from Vu 1982, Figure 30 from N&E 1997, Figure 33 from Hoang 1989).

As we have seen from the spectrograms in Chapter Three and this chapter, *hoi* is a breathy tone and *nga* is creaky. Recall that while both breathiness and creakiness lower the tone, creakiness is produced at very low frequencies and makes a tone lower than does breathiness (Marasek 1997). This is a clear indicator that *hoi* cannot be lower than *nga*: *nga* has creakiness. The fact that *hoi* is not as curved as *nga* also has a physiological explanation: unlike in *nga* where subglottal pressure is built up after the glottal stop and the sudden release of the pressure results in strong turbulence causing high frequency noise, in *hoi* breathiness does not cause high pressure because the vocal folds do not completely close. Therefore, after the breathiness there is not strong air turbulence causing high frequency noise.

Although breathiness is found in *hoi*, I argue now, it is not a distinctive register feature in this tone, but rather it functions to fulfil a requirement for tonal shape. As claimed in the literature, *hoi* is a curved tone (Chapter Three). I proposed in Chapter Two, section 2.2.2, that the curved shape of *hoi* is represented with three points under the contour node. The structure of *hoi* is repeated below in (74). How the phonological feature is phonetically realized will be discussed in the next section.
The pitch graphs in Figure 61 show the consistent curved shape of *hoi* in the forms *[ta], *[ka], *[ma], *[za] and *[sa] from Speaker Dung, in two different tokens. Figure 62 shows a similar pattern from Speaker Khanh in two different tokens, although the tone is less curved than that in Dung, not rising much at the end.
Figure 61. Curved shape of *hoi* in two different tokens from Speaker Dung
Figure 62. Curved shape of *hoi* in two different tokens from Speaker Khanh.
Breathiness in the middle of *hoi* relates to the curved shape of that tone. My study shows cases where the curve is very breathy, as we see in the spectrogram in Figure 63 from Speaker Van.

Figure 63. *Hoi* is very breathy in the middle from Speaker Van

In this figure, *hoi* is in the first and third rows. Breathiness starts at about 90 ms, represented by very weak glottal pulses in the waveform. Breathiness increases to a peak level from about 150 to 200 ms. In the corresponding portion in the spectrogram, there are wide gaps between pulses.

Figure 64 is the corresponding pitch graph.
As we saw in Figure 63, *hoi* is too breathy in the middle for cycles in this part to be measured. However, more typically, breathiness spreads towards the end of the tone *hoi*, and the tone goes down and slightly up shortly before it ends, as seen in Figure 35 from Speaker Son, repeated below as Figure 65. Figure 66 is the corresponding spectrogram of Figure 65.
Figure 65. *Hoi* in [ka] from Speaker Son

Figure 66. Spectrogram of *hoi* in Figure 65 from Speaker Son
In some cases, hoi is breathy throughout the tone and the tone does not go up again toward the end. The curve disappears and the tone flattens toward the end (Figure 32), or even falls as in Speaker Hoang in Figure 30 from Nguyen and Edmondson 1997. When hoi goes down at the end, its contour looks similar to that of huyen, a level tone without creakiness or glottal stop. This phenomenon has also been observed by others in modern Northern dialects, i.e., the tone hoi sometimes loses the rising part (see Figure 30 from Nguyen and Edmondson, other speakers in Figure 32 from Hoang 1989). Vu 1999 says that the loss of the rising part is seen more in female than male speakers. In my study, it is seen equally in both sexes.

When hoi occurs without the rising part, i.e., is breathy and slightly falling, it looks like huyen. How is the missing information recovered to avoid confusion between hoi and huyen? Interestingly, the solution lies in phonation: breathiness has to be stronger in hoi to compensate for the lost curve. Therefore, in order to be distinguished from huyen, hoi is much breathier than huyen, as seen in Figure 67. In this figure, the waveform of hoi is in the second row and its spectrogram is in the last row. In the waveform, between from 70 and 130 ms, the tone is very breathy so that it is difficult to separate individual pulses. The glottal pulses in hoi are much weaker than those in huyen. In the corresponding portion in the spectrogram of hoi, the spaces between cycles are wider than those in huyen. Between 70 and 130 ms where individual pulses are hard to identify, the cycles are irregular.
Figure 67. Spectrograms of *huyen* and *hoi* from Speaker Son

The breathier the tone, the lower it is in F0. As we can expect, in the pitch graph *hoi* is lower than *huyen*. Figure 68 shows the corresponding pitch graph of Figure 67.
In a theory relying solely on pitch height, two different types of 'low' would be needed here: *huyen* is something like 'mid-low' and *hoi* is 'low-low'. A theory that employs only phonation as a way to distinguish tones also needs two types of breathiness here: breathy and very breathy. However, whereas [breathy] is underlyingly specified as a register feature in *huyen*, it is not specified in *hoi*. Instead the breathiness associated with *hoi* is phonetic only.

Moreover, as we have seen, a tonal height-based classification is less preferable than a laryngeal based classification because there are many cases where pitch height does not pattern as expected under the pitch height hypothesis. For listeners who are not very familiar with Northern dialects, or even for listeners of the same dialect, when
hoi loses the rising portion it is easily confused with huyen because phonetically hoi and huyen are similar in this case. Hoang 1986 reports a result from a perceptual test that is relevant here. Participants were asked to recognize tones individually. Only 25% of instances of hoi were correctly recognized; the rest were confused with tone huyen. The fact that hoi can sometimes appear without the rising part might suggest some change in progress in Northern dialects, a neutralization of contour features, i.e., loss of an edge point in the contour feature of hoi. This will be discussed in detail in the next section.

5.1.2.3. Summary and supporting remarks from other studies

In this section I have shown that since F0 varies to the point of not distinguishing a register pair, it is not a reliable cue to differentiate tones. The laryngeal features of breathiness and creakiness are stable and provide a more reliable auditory cues. These features, along with tonal shape, account for tonal height: where breathiness or creakiness occurs, the tone is low. In curved tones, breathiness and creakiness signal where the lowest part is. The features of breathiness and creakiness are found in all marked tones, i.e., huyen, nang1, nga, and nang2. The tone hoi is breathy although it is unmarked for register. The breathiness in this tone satisfies a requirement of the contour feature [curve].

Note that all marked tones except nga are described as low in the traditional literature. Recall that nga is classified as high due to its high ending. In section 5.1.1, I showed that the high ending of nga varies from token to token and can be classified as phonetic rather than distinctive. Therefore, it is difficult to say that phonation predicts
low but not vice versa. However, I can show that at least in the case of nga, creakiness predicts curve, but not vice versa.

All phonological and phonetic properties discussed in this section are summarized in Table 6 below. The rows Contour and Register show phonological properties, except that the phonation of breathiness in hoi is phonetic (in parentheses). The feature [obstruent] occurs only in sac2 and nang2, and distinguishes sac2 from sac1, and nang2 from nang1. The last row shows pitch height, treated here as a phonetic property. For instance, in Table 6 huyen is a level tone, although phonetically it might fall a bit at the end. It has breathy voice that predicts lowness. Nga is a curved tone with creakiness, regardless of how high it ends or how low the middle part is. Nang1 is a falling tone with creakiness, regardless of how high it is before falling due to creakiness, or regardless of how short or long the tone is. Breathiness in hoi is a phonetic feature which is further discussed in section 5.2.

Table 6: Phonological properties of tones (Contour, Register)

<table>
<thead>
<tr>
<th>Tone</th>
<th>nang</th>
<th>huyen</th>
<th>sac1</th>
<th>nang1</th>
<th>hoi</th>
<th>nga</th>
<th>sac2</th>
<th>nang2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contour</td>
<td>level</td>
<td>level</td>
<td>rise</td>
<td>fall</td>
<td>curve</td>
<td>curve</td>
<td>rise</td>
<td>fall</td>
</tr>
<tr>
<td>Register</td>
<td>breathy</td>
<td>creaky</td>
<td>(breathy)</td>
<td>creaky</td>
<td>[obst]</td>
<td>breathy</td>
<td>[obst]</td>
<td></td>
</tr>
<tr>
<td>(Pitch)</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
</tbody>
</table>

Although there has not been any account of Vietnamese that relies on phonation types as phonological features, the important role they play in the tonal system of Asian
languages has been recognized by researchers. Among the studies on acoustic properties of Vietnamese tones, Hoang 1989 agrees that phonation types cannot be ignored in the phonology of Vietnamese tones. He criticizes a common assumption that in Vietnamese, phonological features of the tonal system can be generalized based purely on pitch. If only F0 is taken into account, it is impossible to differentiate, for instance, the tones huyen, hoi, nang1 and nang2. Moreover, he points out how F0 of tone can vary among speakers. According to him, the phonology of the Vietnamese tonal system involves a trade-off between the fundamental frequency and the remnants of segments that still remain during the development of tones from segments, i.e., phonation types. However, he suggests this with no evidence. He does not show how the phonation of laryngeal features can be a distinctive feature, how it is represented, or what the structure of tones is.

Alves 1997 also recognizes that phonation features should be taken seriously as distinctive features. However, he does not give any evidence for this claim or show how those features can be distinctive in the phonology.

Using the phonation types of breathiness and creakiness distinctively is not unusual. Examples of languages that use phonation types distinctively in both consonants and vowels are discussed in detail in Ladefoged and Maddieson 1996. A large number of Mon-Khmer languages such as Phalok, Wa, Chong, Mon, Bru, Kui, So, and Nyah Kur lexically contrast modal voice versus breathy voice (Thongkum 1988). Thongkum reports rare cases where languages have three- or four-way contrasts among phonation types, for instance clear voice versus clear followed by creaky voice, versus breathy voice, versus breathy voice followed by creaky voice (in Chong). In Nyah and
Kui, two Mon-Khmer languages, Thongkum finds that breathy voiced vowels and clear voiced vowels systematically differ from each other only in the laryngeal dimension (phonation types); the rest of the differences (length of the vocal tract and vocal tract shape) are minor. In these two languages, breathy vowels always have lower F0 than clear voice (modal voice) in all types of syllable structures (1988:321).

5.1.3. Summary

In this section I have discussed some interesting implications of the acoustic study for the phonology. I have demonstrated that the mismatch problem between the phonetics and phonology of hoi and nga is not a valid one. Moreover, I have argued that Vietnamese uses phonation distinctively, and it is phonation that predicts tonal height. F0 varies from token to token, but phonation types remain constant. Although there is variation from speaker to speaker in the degree of phonation types (e.g., more or less breathy, glottal stop or creakiness), for each speaker, the presence of breathiness and creakiness in tones with non-modal voice is clearly seen, compared to tones with modal voice. Specifically, creakiness is contrastive in nang1 and nga. Breathiness is contrastive in huyen and nang2. We have seen that the traditional low tone nang1 can be higher than the high tone ngang in most parts of the tone as long as creakiness or glottal stop follows. Without creakiness, nang1 is not perceptible. We have also seen that nga can have a very high or low ending because the middle portion is crucial in this tone. Therefore, without creakiness in the middle of the tone, nga is confused with sac1. The level tone huyen without breathiness would be confused with ngang, another level tone. Nang2 without breathiness is not Vietnamese (there is no such tone). These phonation
types, [creaky] and [breathy], motivate the laryngeal features for tonal register. They are present in all marked tones.

In the following section I will show the relationship between phonetics and phonology of tones in Vietnamese.

5.2. The mapping between the phonological representation and the phonetic realization

5.2.1. Phonation types: motivation for register feature

In (75) I repeat the phonological structure proposed in Chapter Two and discuss the terms used in this representation. Then I show the phonetic representation of tones.

(75) Phonological representation of eight tones

a. \[ \text{ngang} \]

\[ \text{C} \]

\[ \bullet \]

\[ \text{huyen} \]

\[ \text{C} \]

\[ \text{R} \]

\[ \bullet \]

[\text{[laryngeal]}]

[\text{[spread]}]
I review some points on the contour side from Chapter Two. The two level tones *nang* and *huyen* are characterized with one point, indicating that the tone does not change direction throughout its duration. The height of the tone is, as I show below, predictable
from laryngeal features. The rising and falling tones \((\text{sac1}, \text{nang1}, \text{sac2}, \text{nang2})\) are specified with two points, indicating that the tone changes its direction, i.e., up or down or falling-rising. \(\text{sac2}\) and \(\text{nang2}\) have the feature \([\text{obstruent}]\), which indicates that these tones occur only in obstruent-final syllables (see Chapter Six for additional discussion).

We have also seen how important the curve is in \(\text{hoi}\) and \(\text{nga}\). The curve is characterized by three points, which indicates that the tone changes in direction twice, i.e., it goes down and then rises.

As discussed in Chapter Two that I assume \([\text{laryngeal}]\) is a register feature that is underlyingly present in all non-modal voice tones. I use this term to refer to modes of vibration of the glottis in which the vocal folds are further apart or drawn further apart than in modal voice. Following Halle and Clements 1983, I assume that \([\text{constricted glottis}]\) refers to sounds that are produced with the vocal cords drawn together, preventing normal vocal cord vibration. \([\text{spread glottis}]\) refers to sounds that are produced with the vocal cords drawn apart, producing a non-periodic (noise) component in the acoustic signal. From chapter Three and section 5.1.2 in this chapter, we have seen creakiness/glottal stop and breathiness are always found in marked tones to some degree. I use \([\text{constricted}]\) as a distinctive feature for a tone that is produced with creakiness or glottal stop, and \([\text{spread}]\) as a feature for a tone that is produced with breathiness. The phonological feature of register, therefore, is very close to the physiological feature. Both \([\text{constricted}]\) and \([\text{spread}]\) are grouped as features which further specify the feature \([\text{laryngeal}]\).

The phonetic representations of Vietnamese tones are given in (76).
(76) Phonetic representations of Vietnamese tones

a. \textit{ngang} \\
\begin{tabular}{c|c|c}
\textbf{C} & \textbf{C} & \textbf{R} \\
\hline
\textbf{h} & \textbf{l} & \textbf{[spread]} \\
\end{tabular}

b. \textit{sac1} \\
\begin{tabular}{c|c|c}
\textbf{C} & \textbf{C} & \textbf{R} \\
\hline
\textbf{l} & \textbf{h} & \textbf{h} & \textbf{l} & \textbf{[constricted]} \\
\end{tabular}

c. \textit{hoi} \\
\begin{tabular}{c|c|c}
\textbf{C} & \textbf{R} \\
\hline
\textbf{h} & \textbf{l} & \textbf{h} & \textbf{h} & \textbf{[spread]} \\
\textbf{h} & \textbf{l} & \textbf{h} & \textbf{h} & \textbf{[constricted]} \\
\end{tabular}
Comparing (75) and (76), we see that on the contour side all points are realized phonetically, with the value 'h' or 'l' present in the phonetic representations. On the register side, the unmarked tone hoi is realized with the laryngeal feature with [spread] as a dependent. The following section shows how the phonological representations in (75) are implemented as the phonetic representations in (76).

5.2.2. Mapping between phonetics and phonology of tone

In the phonetic representations in (76), all contour features are realized as level, rising, falling or curved. As we have seen in the previous section, in Vietnamese phonation types (i.e., creakiness and breathiness) lower a tone. Tones with non-modal voice have at least a portion of them realized as 'l' by default. This is summarized in (77).
(77) Realization of contour feature

a. A non-modal voice tone is realized as 'l' for Contour on the final point.

b. A modal voice tone is implemented with the feature 'h' on the final point for Contour by default.

Let us look at (75) and (76) to see in detail how contour features are realized and what the possible variations are. With the two level tones ngang and huyen, they both have one point specified for the contour feature in (75). Note that neither ngang or huyen is underlyingly specified for the value 'h' or 'l' under the contour node. Ngang with modal voice is realized phonetically as 'h'. Huyen with the register feature [spread] is realized as 'l'. This, I propose, is straightforwardly the result of the default rule in (77).

Sac1 is a rising tone, therefore it is specified with two points to indicate the change in direction. Recall the assumption in Chapter Two, section 2.2.2, that two points at the phonological level must be interpreted as having different values. Recall also that the end point is most crucial in the recognition of a tone. Since sac1 has modal voice, the endpoint is realized as 'h'; the onset has the value 'l', generating the contour 'lh'. The rising part might come early or late in the tone. It does not matter how late the tone rises, as we have seen in Chapter Three, section 3.3.3. As long as sac1 rises a little at the end, it is not confused with ngang, a high level tone. The late rising part in sac1 is a typical pattern in Hanoi speakers. Similarly, sac2 is a rising tone with modal voice. The contour, therefore, is 'lh'. Sac2 is very short; the tone has to rise early before it can end.

Nang1 is a falling tone. Nang1 has creakiness captured by the register feature [constricted]. Because of the register feature, the end point is realized as 'l'. The first
point is realized as 'h', a different value from that of the first point, generating the falling contour ‘hl’. We also have seen that it is not important whether the tone is 'high' or 'low' before it is pulled down by creakiness or glottal stop. Because of the big perceptual difference between creakiness and breathiness, one would expect that a perception experiment would show that as long as there is creakiness in the tone, that tone is not confused with the low level tone *huyen* with breathiness. Similarly, *nang2* has breathiness, represented by the register feature [spread] which lowers the tone. The end point is realized as ‘l’ from [spread], generating the falling contour ‘hl’. Without the feature [obstruent] that shortens the tone, *nang2* might be identical with *huyen* in pitch height as we saw in Chapter Three.

Now let us look at the two curved tones, *hoi* and *nga*. We have seen how important the curve is in *hoi* and *nga*. The tone changes direction twice to make the curve, therefore, it is specified with three points. Recall my assumption from Chapter Two, section 2.2.2 that if a tone is specified for more than one point, the adjacent points have different values. For instance, if there are two points specified, the contour has to be either ‘x y’ or ‘y x’, e.g. as in *sac1* or *nang1*, but not *‘x x’* or *‘y y’*. If there are three points specified, the first and last points have a different value from the middle point. This assumption generates the contour ‘hlh’ in *hoi* and *nga*. Here we might ask a question: why are *hoi* and *nga* not ‘hl’? Recall from Chapter Three, section 3.3.1.2 that curved tones in Vietnamese are concave, i.e., falling-rising. I assume that the contour *‘hlh’* is not allowed in Vietnamese; therefore ‘hlh’ is the only possible contour. This contour shape forces the tone to go down in the middle. Let us first look at the *nga* tone. *Nga* has the register feature [constricted], which is motivated by creakiness or glottal stop in the
middle of the tone. The only aspect of this tone that is odd is that the contour is ‘hlh’; the constraint against *hlh must take precedence over the general rule that the non-modal tone has a final ‘I’. Returning to the contour of nga, we also saw in section 5.1.2.2 that the degree of creakiness can vary from speaker to speaker, but it always occurs to make a curve, even a slight curve, as in Figures 56 and 57. Note that in nga, the rising part after creakiness is crucial. In nang1, we see that sometimes the tone rises a little after creakiness. However, this rising part is not crucial in nang1 because nang1 is a falling tone. Therefore, we see variation in nang1: sometimes nang1 ends suddenly after glottal stop; sometimes it goes further, flattening out or rising a bit. Consequently, whether the starting point in nga or nang1 is very high in the speaker’s pitch range, or quite low, the tone is equally perceptible if there is creakiness in the middle of tone nga to make the curve, or in tone nang1 without making a curve.

I now turn to the curved tone hoi. Like nga, I specify hoi with three points for contour, without a laryngeal feature underlingingly. The reason to the contour ‘hlh’ is clear: given that like tones cannot occur adjacent to one another and that non-modal voice is present phonologically, ‘hlh’ is predicted. We must now ask why this tone has breathiness phonetically. I propose that it is the low mid point of this tone which forces it to be breathy. Recall that both creakiness and breathiness are compatible with low F0. Because the contour is curved with three specified points, this contour predicts either creakiness or breathiness in the middle. We have seen that creakiness is phonologically a laryngeal feature in nga, I assume that breathiness is chosen as a default feature in hoi. Because the curved contour is phonological in hoi, I assume that this tone gets breathiness by default in order to satisfy the requirement from the contour side: the
curve requires a laryngeal feature in the middle of the tone. Breathiness, therefore, is a phonological feature in huyen and nang2, but a phonetic feature in hoi.

Then a question arises: why is the feature [laryngeal] realized as [spread] but not [constricted] in hoi? I assume that spread glottis is less marked than constricted glottis. With glottalization, the vocal folds have to make a complete closure for a glottal stop, or they have to be tense enough to produce creaky voice. In breathy voice, the vocal folds are lax to open. Therefore, it takes less effort to produce breathy voice than creaky voice.

Assuming the Gestural Theory of Markedness (Hamilton, 1996:8), articulatory gestures which deviate minimally from configurations of least effort are highly valued. With respect to the frequency of occurrence of unmarked features, Greenberg 1966 shows evidence from a diverse sample of languages that non-glottalised consonants, including aspirated consonants, occur more frequently than their glottalized counterparts. This is expected if we assume glottalized consonants involve more effort than non-glottalized consonants.

I summarize the realization of contour features in (78) below, modified (77).

(78) Phonetic realization of tones

a. A non-modal voice tone is realized as ‘l’ for Contour on the final point.

b. A modal voice tone is implemented with ‘h’ on the final point for Contour by default.

c. hoi gets [spread] for Register by default to fulfil a requirement for the curved contour.
5.2.3. Neutralization in hoi

5.2.3.1. Neutralization of Contour feature in hoi: hoi loses the rising part

Recall that for many Hanoi speakers, hoi is realized as a falling tone without the curve. In this case, it is easily confused with huyen. Although this variation is not seen in other Northern dialects, it is quite common in the Hanoi dialect. This can be seen as a neutralization of the contour feature. The neutralization process is shown in (79). Note that (79) shows a phonetic representation.

(79) Neutralization of contour feature in hoi

\[
\begin{align*}
\text{(a)} & \quad \text{hoi} & \quad \text{(b)} & \quad \text{hoi} \\
\text{C} & \quad \text{R} & \quad \rightarrow & \quad \text{C} & \quad \text{R} \\
\bullet & \quad \bullet & \quad [\text{laryngeal}] & \quad \bullet & \quad \bullet & \quad [\text{laryngeal}] \\
\text{h} & \quad \text{l} & \quad \text{h} & \quad [\text{spread}] & \quad \text{h} & \quad \text{l} & \quad [\text{spread}] \\
\end{align*}
\]

The last point under the contour node, which constitutes the curved shape for the contour in this tone, is delinked. There are only two points left for the contour and breathiness under the register lowers the tone to create the falling contour ‘hl’. On the phonetic level, the result makes hoi look very much like huyen, sharing its breathiness but with a falling contour. Note that in (79b), the simplified hoi, the contour is ‘hl’ while in huyen the contour is ‘ll’ phonetically. However, as mentioned above, the contour node in huyen is not underlyingly specified for ‘l’ or ‘h’. Assuming that lack of specification
can lead to some variation (e.g., Archangeli 1988, Pulleyblank 1986, 1988, Rice 1996): 
*huyen* might be realized as a level tone, or might go down a little at the end as a falling tone (e.g., Nguyen and Edmondson 1997 classify *huyen* as a 21 tone, compared to 44 for *ngang*). This explains why *hoi* is confused with *huyen* when the rising part is neutralized, although, as discussed in section 5.1.2.2, in order to be distinguished from *huyen*, Hanoi speakers increase the breathiness in *hoi* after it loses the rising part. The result is that *hoi* is much breathier than *huyen*.

There is another neutralization pattern of *hoi* in other dialects. Vuong and Hoang (1994:103) observe that in some dialects of Quang Binh province in Central Vietnam, *nga* neutralizes to *nang1* and *hoi* neutralizes to *sac1*. The case of *hoi* neutralizing to *sac1* can be accounted for as follows. In this process, *hoi* loses the falling part, and only the rising part remains, resulting in *sac1* with the rising contour as shown in (80). This pattern is similar to one in the Mai Ban dialect discussed in Chapter Two, section 2.3.5.2. (80) shows the neutralization process on the phonological level.

(80) Neutralization of contour of *hoi* in Quang Binh dialects

\[
\begin{array}{c}
\text{a. } \text{hoi} \\
| \\
\text{C} \\
\text{↑} \\
\text{• • •} \\
\end{array}
\quad \rightarrow \quad
\begin{array}{c}
\text{b. } \text{sac1} \\
| \\
\text{C} \\
\text{→} \\
\text{• • •} \\
\end{array}
\]

Note that the issues of the input for neutralization - phonetic representation or phonological representation - are important, but I do not deal with them here.

247
5.2.3.2. Neutralization of register feature in *hoi* in Southern dialects

In Chapter Two I discussed some neutralization phenomena in other dialects of Vietnamese. For instance, in Southern dialects, the tones *hoi* and *nga* are merged phonologically and realized as *hoi*. How does this process happen phonetically?

Recall that the proposed analysis suggests that *nga* is structurally more complex than *hoi* and that phonation, not pitch height, is distinctive in Vietnamese tones. Greenberg (1966: 16-34), among others, observes that when the contrast between two segments is neutralized in a particular context, the neutral segment which occurs is often phonetically closest to the unmarked member of the set. Neutralization is in favor of unmarked features. This predicts that in the neutralization of *nga*, the distinctive feature in *nga* is lost, resulting in some tone that is similar to *hoi*, the unmarked member of the set. We saw that delinking of register in *nga* results in *hoi*. What is interesting here is the phonetic contour of the resulting tone, which survives neutralization. This contour is phonetically half-*hoi*, half-*nga*. Let us see how this happens.

When the two tones *hoi* and *nga* merge in Southern dialects, one tone disappears. The resulting tone has a very interesting contour feature. Recall that *hoi* and *nga* share the same contour feature - they are both curved. The register feature [constricted] in *nga* is delinked, resulting in the new tone, which is labelled 'hoi'.
(81) Neutralization of *hoi* and *nga* in Southern dialects

With a curved shape, the new tone looks like the Northern *hoi* and it is recognized as *hoi* for most speakers in both Northern and Southern dialects. However, the new tone, perhaps in order to compensate for the lost register feature, rises up quite high and clearly occupies the full range of pitch, i.e., between maximal and minimal F0 values. With this shape the new tone cuts across both the lower and higher parts of the pitch range. Seven tones in the Saigon dialect from Hoang 1989 are shown in Figure 69, repeating Figure 27 from Chapter Four.
In Figure 69, *hoi* has the same shape in all four speakers. Because of this shape, some authors suggest labelling the new tone as 'hoi-nga' (Hoang T. Chau 1989). We also find the same shape of *hoi* in Southern dialects from Vu 1982, given in Figure 70.
This variation in the resulting hoi in Southern dialects shows that, being unmarked for the register feature, the contour feature of hoi can vary to make a very high ending as a trade-off for the marked laryngeal feature lost in nga.

In summary, in this section I have shown how the phonological features of tone are motivated and realized, and how they are implemented. The contour feature is realized as 'T' if the tone has a marked register feature, as in huyen, nang1, nang2 and nga. Phonation types of creakiness and breathiness in these tones generate a low contour in level and falling tones, or in the middle of the curved tone nga. Otherwise, the contour is realized as 'h' in tones with modal voice as in ngang, sac1 and sac2. The tone hoi is special
in that, as an unmarked register tone, it gets breathiness by default in order to satisfy the contour requirement. Thus, the realization of tone is generated from the configuration of both tonal shape (contour) and phonation types (register). I have also shown how hoi is neutralized and realized phonetically in other dialects.

5.3. Summary

In this chapter I have provided evidence from my instrumental study to argue that the phonation types creakiness and breathiness are reliable in Vietnamese and can group tones into natural classes. Instead of pitch height, the laryngeal features of creakiness and breathiness are distinctive as the register feature in Vietnamese tones. Tonal height is derived from the configuration of contour and register, or tonal shape and phonation types, respectively. Using phonation types distinctively not only can explain the various tonal patterns in the language, it also makes the problematic mismatch an illusory one, which in turn, makes ad hoc, unnatural flip-flop rules unnecessary. The relationship between phonetics and phonology, therefore, shows phonological features and physical gestures are very close in Vietnamese.

It is reasonable to assume that in Vietnamese the feature Register originated historically from the loss of initial voicing (i.e., losing voiced initials creates low register while the original voiceless initials give high register). However, when both voicing and tone became contrastive, there is no reason to maintain the assumption that register is tonal height.
Chapter 6

DOMAIN OF TONE

In this chapter I discuss the domain of Vietnamese tone, arguing that the domain is bigger than the nucleus but smaller than the syllable, based on both phonetic and phonological evidence. The domain of tone is the rhyme. Phonological evidence comes from the distribution of short and long vowels in the rhyme, and the intimate relationship between the vowel and final consonant through feature sharing. Phonetic evidence comes from the distribution of the tonal features and from equal rhymal length with different vowel types. In Section 6.1 I discuss the structure of the syllable in which tone is independent from segments, and the constituents in a Vietnamese syllable which can be candidates for the domain of tone. In section 6.2 I examine all possibilities of a tonal domain including the syllable, rhyme, nucleus, and mora. The conclusion is in section 6.3.

6.1. Syllable structure in Vietnamese

In this section I show that tone is independent from segments, and I provide evidence for each constituent in the syllable structure.
6.1.1. Tone is independent from segment

As we saw in Chapter One, the Vietnamese syllable has four components: tone, initial consonant, vowel and final consonant. The initial or final consonant can be absent while the vowel and tone are required. There is evidence from reduplication, poetry and language games that tone is independent of segments.

First, as seen in Chapter Two, section 2.3, there are several cases in which tone patterns independently of segmental features. For instance, in total reduplication, whole syllables are kept but tones are changed. In partial reduplication, either the initial consonant or rhyme is kept, but the tone is changed. Second, in poetry (also from Chapter Two), syllables in the rhymal position (where only words having the same rhyme can occur) keep a constant rhyme, but the initial consonant and tone can be changed. Finally, a third source of evidence is often cited in the Vietnamese literature (Doan 1977, Vuong and Hoang 1994). Many researchers identify a language game in Vietnamese which can switch around onsets, tones, and rhymes within a morphologically complex word (Doan 1977, Vuong and Hoang 1994). This game is illustrated in (82), taken from a beautiful anecdote in folk literature, using words with non-stop finals. Nonsense words do not have glosses in (82).

(82) Word game

lo (nang1) tuong (ngang)  >  a. tuong (nang1) lo (ngang)  ‘a worried statue’
‘a bottle of sauce’       >  b. luong (nang1) to (ngang)  ‘great merit’
                      >  c. luong (ngang) to (nang1)
                      >  d. to (nang1) luong (ngang)
In (82), in the original syllables 'lo tuong' (a bottle of sauce), the first syllable bears the tone nang1, the second syllable bears the tone ngang. This word can occur in five different new ways. In (82) all possible combinations are given although (82c) is a non-sense word. In the derived form in (82a) the two original syllables, lo (nang1) tuong (ngang), exchange their initials and rhymes, keeping their tones: 'lo' with the original tone nang1 becomes 'tuong' keeping nang1, 'tuong' with the original tone ngang now becomes 'lo' keeping tone ngang. In (82b) the two syllables exchange non-tonal rhymes only: e.g., 'lo' keeps the initial 'l' and tone nang1 but has the rhyme 'uong'. In (82c) they exchange both tone and rhyme: e.g., 'lo' keeps the initial 'l' but has the tone ngang and the rhyme 'uong'. The form (82a) derived from (82) occurs in a well-known folktale in Vietnamese literature. Note that the vowel can switch independently of tone (82a, b) and vice versa (82a, b, e). The ability to be detached freely from the original syllable of the initial, rhyme and tone as seen in (82) provides an argument for the independence of these elements in the Vietnamese syllable, and for tone being independent from segments (Doan 1977).

All these facts show that tone is independent from segments. The independence of tone is recognized in autosegmental theory by putting tone on an independent tier (e.g., Leben 1973, Williams 1976, Goldsmith 1976).

In the traditional Vietnamese literature, there is a view (e.g., Le 1948, Hoang, Le and Cu 1962) that does not count tone as a constituent of syllable structure, because tone is not a segment, and thus is not regarded as a phoneme. Other researchers emphasize
the function of each component in the syllable (Mkhitarian 1959, cited in Doan 1977), but it was not until Gordina 1959 (cited from Doan 1977), that the structure of the Vietnamese syllable was recognized as a phonological unit. Although Gordina did not include tone in the syllable structure, according to her, all components in a syllable relate to each other in a hierarchical relationship. All researchers share the idea that tone is on a different level from the level of segments. For instance, Doan (1977:88) proposes that the syllable has the structure in (83) in which tone is on a separate level.

(83) Structure of Vietnamese syllable

<table>
<thead>
<tr>
<th>Initial</th>
<th>Tone</th>
<th>Rhyme</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>/w/ 21</td>
<td>Vowel</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Others adopt similar structure.

In summary, the fact that tone can be separated from segments in several patterns in the language shows the independence of tone from segments. Before discussing the issue of what tone associates to, in the following section I provide evidence for the constituents of the syllable.

---

21 The labio-glide /w/ is the only sound that can occur between the initial consonant and the vowel. The /w/ is regarded either as a labial feature of the initial consonant ((Le 1948, Hoang et al 1962) or as an independent segment which belongs to the rhyme (Doan 1977:87). Also there is one process of reduplication in which the rhyme is always [tvk] plus the initial consonant of the base. If the base has /w/, usually only the /w/ is not reduplicated in the reduplicant. For example, [tva:m] ‘mathematics’ becomes [tva:m tvk] (Doan 1977). Based on such words, /w/ is argued to be independent from the initial. Turning to the former view, the labio-glide /w/ in this position is treated as rounding of the initial consonant, because /Cw/ is the only consonant cluster in Vietnamese in this position. In this view, the initial /t/ in /t^a:m/ ‘mathematics’ has the rounding feature. I am not concerned with whether /w/ is a segment or a feature of the initial.
6.1.2. Independence of initials: evidence for rhyme

In this section, putting aside tone, I show that the initial forms one constituent while the rhyme forms another. Evidence for the independence of the initial comes from reduplication and a language game, and it shows that the rhyme is a possible tone-bearing unit.

Consider first the language game discussed above in section 6.1.1. In this game the initials of two syllables can be switched (as in (82c, d)). This shows that the onset is independent from the rhyme. In reduplication, there are several patterns in which either the initial or the rhyme is reduplicated. Some data are given in (84). The base is underlined.

(84) Reduplication: evidence for the initial being independent from rhyme

a.  laːn  ‘cold’  >  laːn  lew  ‘very cold’
    sin  ‘pretty’  >  sin  san  ‘very pretty’

b.  duː  ‘exhausted’  >  luː  duː  ‘very exhausted, motionless’
    kurŋ  ‘fall into panic’  >  lurŋ  kurŋ  ‘fall into panic, get confused’

In (84a) the initial is reduplicated, in (84b) the rhyme is reduplicated. The rhyme is unpredictable in this process.\(^{22}\) There are other processes of reduplication in which only the initial remains. For instance, there is a process in which the reduplicant is a

\(^{22}\) The process, perhaps, was productive once, but is not any longer in modern Vietnamese.
combination of the initial of the base and the rhyme [irk], as seen in (85). The meaning is derogatory (e.g., Doan 1977, Ngo 1984).

(85)  ᵇᵒᵏ  'to study'  >  ᵇᵒᵏ  ᵇⁱʳᵏ  
      ᵇᵃⁿ  'table'        >  ᵇᵃⁿ  ᵇⁱʳᵏ

The process in (85) is very productive. These facts show that the initial can be easily separated from the rhyme.

There is also a process in reduplication that reduplicates the initial, final and tone but not the vowel, e.g., [lak] sac2 'to shake' > [luk] [lak] sac2 - sac2 'to wag'. This process is not productive, but speaks to the independence of the nucleus. In addition, rhymal syllables in poetry, discussed in section 6.1.1, also support the claim that the relationship between the vowel and the final consonant is much closer than that between the initial and the rhyme.

In the traditional literature, there are two views concerning the internal structure of the syllable. In the first view (Le 1948, Hoang, Le and Cu 1962), the initial, the vowel and the final have an equal and independent status in the syllable, represented by the plus signs: C₁ + V + C₂. In the second view, a syllable is represented as a unit with constituents. Gordina (1959) identifies a close relationship between the vowel and the final consonant based on her experimental phonetics work. She points out that phonetically the rhyme in Vietnamese syllables occupies a constant amount of time regardless of segmental content. By recognizing the relative independence of the initial consonant from the rhyme and the close relationship between the vowel and final

258
consonant, researchers (Doan 1977, Vuong and Hoang 1994, Cu, Hoang and Nguyen 1977) divide a Vietnamese syllable into two major components: the initial and the rhyme. This can also be seen in (83) above, and in (86) below from Vuong and Hoang (1994:78).

(86) Structure of the syllable

<table>
<thead>
<tr>
<th>Syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

In (86), the syllable has three major components: tone, initial and rhyme. The vowel and final consonant are dependents of the rhyme. The following section will look at constituents inside the rhyme.

6.1.3. Feature sharing: nucleus vs rhyme

In this section I discuss evidence for a distinct nucleus and coda in Vietnamese. There are some phonological processes that take place in the nucleus but not in the whole rhyme. This fact gives evidence that the rhyme and nucleus are two different domains in Vietnamese. Assuming, as discussed above, that the nucleus and final consonant form a constituent, then the nucleus is a constituent of the rhyme.
Components of the nucleus relate to each other in a very special way. In all dialects of Vietnamese, the realization of final velar consonants depends on the quantity and quality of the vowel. Of interest here is vowel quantity.

Recall that only a limited number of consonants can occur in final position. The phonetic final inventory of the Hanoi dialect given in Chapter One is repeated here as (87). I focus on the final consonants; the glides /j/ and /w/ are omitted since they behave differently from the final consonants (see Pham 1997a). (87a) shows the phonological final inventory (see Pham 1998). (87b) shows the phonetic final inventory.

(87)

a. The phonological final inventory

<table>
<thead>
<tr>
<th>labial</th>
<th>coronal</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>t</td>
<td>k</td>
</tr>
<tr>
<td>m</td>
<td>n</td>
<td>ɲ</td>
</tr>
</tbody>
</table>

b. The phonetic final inventory in the Hanoi dialect

<table>
<thead>
<tr>
<th>labial</th>
<th>coronal</th>
<th>palatal</th>
<th>velar</th>
<th>labio-velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>t</td>
<td>c</td>
<td>k</td>
<td>kp</td>
</tr>
<tr>
<td>m</td>
<td>n</td>
<td>ɲ</td>
<td>ɲ</td>
<td>ɲm</td>
</tr>
</tbody>
</table>

In Vietnamese, the quantity and quality of the vowel determine the surface form of final velar consonants. For instance, in the Northern dialects, if the vowel is short, the final velar consonant is realized as a labio-velar consonant after round vowels, and as a palatal consonant after front vowels. It is realized as a plain velar consonant after central
vowels, long vowels, and diphthongs (Pham 1998). The surface distribution of the final velar consonants is shown in (88). A nasal is illustrated, but the consonant can also be an obstruent. After short front vowels, only [ɲ] occurs. After short back vowels, only [ŋm] occurs. After central vowels, long vowels or diphthongs, only [ŋ] occurs. In (88) V stands for a short vowel, V: stands for a long vowel.

(88) Distribution of final velar consonants in the Hanoi dialect

<table>
<thead>
<tr>
<th>VC</th>
<th>V:C</th>
</tr>
</thead>
<tbody>
<tr>
<td>front</td>
<td>central</td>
</tr>
<tr>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>e</td>
<td>o</td>
</tr>
<tr>
<td>ë</td>
<td>ë:</td>
</tr>
<tr>
<td>n</td>
<td>ɠ</td>
</tr>
</tbody>
</table>

The realization of a final velar consonant, then, depends first on vowel length and second on which feature it receives from the vowel. The consonant shares the feature of the preceding front or back vowel if the vowel is short. The central vowel does not have any feature to share (see Pham 1998). If the vowel is long, feature sharing between the vowel and final consonant is not found.

Assume that a short vowel contributes one mora and a long vowel or a diphthong contributes two moras. If we assume also that the nucleus is minimally and maximally bimoraic, we can say that the consonant is affected by the vowel only if it is moraic, or in the nucleus domain, i.e., when the vowel is short. It is not affected by the vowel if it is non-moraic, or outside of the nucleus domain, i.e., when the vowel is long.
or a diphthong. Only nuclear consonants are subject to feature sharing with the vowel. This is shown in (89). \( F \) is some feature.

\[(89) \text{ Bimoraic nucleus: domain of feature sharing in Vietnamese rhyme.}\]

\[
\begin{array}{c}
\text{a. VC} \\
\text{Ri} \\
\text{Nu} \\
\text{V} \\
\text{F}
\end{array}
\quad
\begin{array}{c}
\text{b. V:C} \\
\text{Ri} \\
\text{Nu} \\
\text{V} \\
\text{F}
\end{array}
\]

\( V \) can affect \( C \)  
\( C \) can not be affected by nuclear Vs

In (89a), the vowel is short, the final consonant is inside the nucleus and receives a place feature from the vowel. In (89b), the vowel is long, and the final consonant is outside of the nucleus domain. It thus cannot receive a place feature from the vowel. Interested readers can see Pham 1998 for a detailed account.

Another phenomenon is also observed in the rhyme if the vowel is short: the vowel is centralized. For instance, in the Saigon dialect, short front vowels surface as a central vowel of the same height before \([n]\) or \([t]\), i.e., /i/ surfaces as [u]; /e/ surfaces as [y]; /ɛ/ surfaces as [a] (see Pham 1998 for details).

---

23 Whether the final consonant in Vietnamese is moraic or not does not affect the argument here.
In summary, I have shown that the relationship between a short vowel and a following consonant is a very close one. Assuming that VC forms a nucleus in the Vietnamese rhyme, then the fact that feature sharing and centralization of vowels do not occur if the vowel is long shows that the nucleus and the rhyme are not the same domain for certain phonological processes. The nucleus obligatorily contains two moras, with the second one filled by either a vowel or a consonant. Further material is in the rhyme. The processes illustrated here have the nucleus as their domain.

I now turn to the issue of which of these constituents, the nucleus or the rhyme, tone takes as its domain. I will show that it is not the syllable as a whole, or the mora or the nucleus, but something bigger than the nucleus and smaller than the syllable, namely the rhyme.

6.2. Domain of tone: syllable, rhyme, nucleus, mora

In the literature on tone, there are different views on what the tone-bearing unit is. Goldsmith 1976 suggests it is the vowel, while Clements and Ford 1979 argue for the syllable as the tone-bearing unit. See Odden 1995 for an overview. Bao 1999b, based on Chinese dialects, argues that the rhyme is the tone-bearing unit. In this section I examine each of the units in the Vietnamese syllable. I argue that the tone-bearing unit is the rhyme.

The following section provides evidence that the tonal domain is smaller than the syllable. We have seen that phonologically, the initial is independent from the rhyme (section 6.2.1). I will show that phonetically tone is also independent of the initial.
6.2.1. The syllable is not a domain: Initials are outside of the domain

In section 6.1.2, we saw that the initial consonant is independent of the rhyme. In this section I provide evidence from acoustics to show that the initial is outside of the domain of tone.

First I briefly examine the literature concerning the tone-bearing unit. There are two views. The first (e.g., Cu et al 1972, 1977) puts tone as a component that associates with the rhyme, as in (90). This view is based primarily on the acoustic results from Andreev and Gordina 1957 which show that tone is only realized on the rhyme, and the initial does not contribute to the value of tones.

(90) Structure of Vietnamese syllable

<table>
<thead>
<tr>
<th></th>
<th>Tone</th>
<th>Rhyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>vowel</td>
<td>final</td>
</tr>
</tbody>
</table>

The second view (e.g., Doan 1977) puts tone over the whole syllable, as seen in (83). Huu and Vuong 1980 share this view, but note that the typical contour of each tone lies on the rhyme (1980:66).

Since tones do not start until late in the vowel (see following sections and Chapter Three), it appears that tone has no effect on the onset. For instance, one might expect tones to affect laryngeal features of onsets, but they do not: any onset can occur with any tone. Bao (1999b:10) shows that tones do not affect onsets in some Chinese dialects, and Vietnamese falls in this category.
If the domain of tone is not the syllable as a whole, other possible candidates now are the mora, the nucleus and the rhyme. I will examine each in the following sections.

6.2.2. Mora is not the domain of tone

Evidence that the tonal domain is not the mora comes from vowel quantity and tonal distribution.

All tones can occur with both short and long vowels in both syllable types, as shown in (91).

(91) Distribution of tone in syllable types

<table>
<thead>
<tr>
<th>long vowel</th>
<th>short vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. sa:w (ngang)</td>
<td>saw (ngang)</td>
</tr>
<tr>
<td>b. sy:n (ngang)</td>
<td>sy:n (ngang)</td>
</tr>
<tr>
<td>c. da:ŋ (sac1)</td>
<td>da:ŋ (sac1)</td>
</tr>
<tr>
<td>d. rr:t (sac2)</td>
<td>rr:t (sac2)</td>
</tr>
<tr>
<td>e. l:r:t (nang2)</td>
<td>l:r:t (nang2)</td>
</tr>
<tr>
<td>f. fa:j (hoi)</td>
<td>fa:j (hoi)</td>
</tr>
<tr>
<td>g. ha:j (nga)</td>
<td>ha:j (nga)</td>
</tr>
</tbody>
</table>

Syllables in the first column have long vowels, forming minimal pairs with those in the same row in the second column. These forms differ from their long vowel counterparts only in that the vowel is short. A moraic analysis for the placement of tone where the
tone occupies only a single mora cannot account for the fact that tones can take either the long or the short vowel as their domain.

6.2.3. The nucleus is not the domain of tone

The fact that a consonant must be in the nucleus in order to share features with a short vowel suggests that the nucleus is a possible domain of tone.

In this section, I provide acoustic evidence to argue that the nucleus is not the domain of tone. First, the tonal feature is distributed over the whole rhyme, not just on the nucleus. Second, non-obstruent final syllables tend to be equal in length, which means their rhymes share the same length.

6.2.3.1. Tone is distributed over the whole rhyme

The acoustic study shows that the tonal feature is not always located at a particular point within the vowel, e.g., onset, middle point or end point, but rather it is distributed over the entire rhyme. Moreover, in sonorant-final syllables, the location of the distinctive feature of a tone within the rhyme depends on vowel length. I show below the distribution of the tones nga and sac1 in the rhyme.

In the Vietnamese literature, based on the acoustic studies of Andreev and Gordina 1957, authors such as Doan 1977 and Vuong & Hoang 1989 describe tone as distributed over the rhyme, with the distribution of tonal features varying depending on rhyme type. For instance, in non-obstruent final syllables, Doan (1977:111) describes the distribution of the tone nga as follows: the tone goes down suddenly and rises up, or it is broken in the middle. In an open syllable, the lowest or broken portion usually falls at
the beginning of the second part of the rhyme. In a syllable with a final nasal, this portion can fall on the final nasal, especially if the vowel is short. A few points are worthy of mention. First, as seen in Chapter Three, creakiness or glottal stop in the middle of the curved tone nga is a distinctive feature of this tone. It can be located anywhere in the tone except for the beginning and ending as long as the tone has a curved shape. Assuming that the broken portion usually falls on the middle of the tone, Doan notes that in a final-sonorant syllable, the broken part falls on the final consonant, which shows that tone is spread over the whole rhyme of at least a sonorant-final syllable, not just over the vowel. Secondly, Doan emphasizes that when the vowel is short, the broken portion falls on a final nasal. This point shows that tone is sensitive to syllable length and adjusts the location of its features in order to be realized over the rhyme. Doan also gives a similar distribution of the tone hoi over the rhyme, i.e., the lowest part of this tone falls on the middle of the rhyme, and on the final sonorant if it is a closed syllable.

My experimental study gives a similar picture. For instance, in Chapter Three, we saw that nga is a curved tone. It has either a glottal stop or creakiness in the lowest part of the tone. The part with glottal stop or creakiness is usually in the middle of the vowel, but sometimes it can be distributed anywhere from right after the beginning of a vowel and before the end of a final sonorant. Figure 71 shows the tone nga in an open syllable.
In Figure 71, because the syllable ends with a vowel, the broken portion (a wide gap in the spectrogram) is in the middle of the vowel, from approximately 100 to 120 ms.

Figure 72 shows nga in a nasal-final syllable with a long vowel [tə:n] and a short vowel [tam], respectively. The arrows show approximately beginnings and end points.
In Figure 72, where the syllable ends with a sonorant consonant, the nasal part is represented with very little voicing in the spectrogram. Here the broken part is not in the middle portion of a vowel as it is in an open syllable, but in the middle of the rhyme, which is the last part of the vowel and the beginning part of the final nasal consonant.

The broken part (or creakiness) is from approximately 100 to 150 ms into the syllable with a long vowel (the first and third rows). In a syllable with a short vowel (the second and fourth rows), this part, represented by the gap in the spectrogram, is much longer, and comes much earlier, from where the vowel ends (around 50 ms) to the beginning of the nasal (150 ms). Recall that in open syllables (Figure 71) the laryngeal feature in nga ends before the end of the vowel. In nasal-final syllables (Figure 72) the laryngeal
feature in nga goes to the end of the vowel. The spectrogram in Figure 72 does not clearly show whether creakiness in nga goes into the final nasal, but this is not a crucial point because a glottal stop or creakiness occurs in the middle of the tone. The important point is that the laryngeal feature does not occur in the middle of the vowel, when the vowel is long, in nasal-final syllables. The fact that the tonal feature is not always in the middle of the vowel, but moves toward the final consonant, supports the claim that tone is not a feature of the vowel alone, and that it is realized over the entire rhyme.

What happens if the distinctive feature of the tone is at the end point of a vowel? In this case, we see a similar pattern: regardless of rhyme type, a tone always distributes its feature over the whole rhyme. For instance, recall that sac1 is a rising tone. It goes flat and rises up. The rising part is distinctive in sac1 and makes it different from, for example, nang1, a falling tone that goes down due to the creakiness. The rising part of sac1 is usually in the second half of the rhyme. In some speakers, however, it sometimes can come very late, just before the rhyme ends. According to Doan, in nasal-final syllables, the rising part occupies the second half of the rhyme, i.e., the final sonorant. Figure 73 shows the pitch graph of the tone sac1 from Speaker An, with and without the final nasal consonant. We see that in both the open syllable [ta] and the closed syllable [taːŋ], the tone has a similar fundamental frequencies through the whole rhyme: the flat portions lie on top of each other, and strikingly, in both syllables, the tone starts to rise at approximately 150 ms.
Figure 73. Tone $sac_1$ in [ta] and [ta:ŋ] from Speaker An.

The corresponding spectrogram of Figure 73 is shown in Figure 74.
From Figure 74, we see that the final nasal starts around after 150 ms. Thus, in [ta] the rising part is in the second part of the vowel, but in [ta: η], it does not come until the beginning of the nasal.

More interestingly, if we look at the tone sac1 in the two syllables [ta: η] and [tan] in Figure 75, we see that the rising part falls in the same location in both syllables, i.e., the final consonant. In the diagram, 'tang' stands for [ta: η], 'ta(ngg)' for [tan].
In the pitch graph in Figure 75, the rising part in both syllables starts after 150 ms. The corresponding waveforms of the tone sac in both syllables [ta:n] and [taŋ] is shown in Figure 76. In the syllable with a long vowel, i.e., the third row of the waveform, the rising part starts just before the end of the vowel (around 150 ms). In the syllable with a short vowel in the last row, the vowel ends at approximately 90 ms. The rising part did not start from the end of the vowel, but from the second half of the final consonant (around 150 ms). This fact shows that it is not important which component in the rhyme the tone links to, but the tonal feature is distributed flexibly over the whole rhyme.
Figure 76. Spectrograms of [taŋ] and [tan] with the tone sac1 from Speaker An

From the facts above, we can predict that in sonorant-final syllables, if the final consonant were omitted, the distinctive feature would be not clearly present, making the tone unrecognizable. Long vowels are especially interesting, as in V:C too the tones is on the final consonant, evidence that the domain of tone goes beyond the nucleus. Although perception tests are not carried in this study, this prediction is likely to be correct. From Figure 76, we see that if the nasal consonant is cut off, the pitch graph shows only the flat part, and the tone is predicted to be perceived as the level tone ngang.
Obstruent-final syllables show a different picture concerning the distribution of tonal features over the rhyme. Recall that only sac2 and nang2 can occur in this type of syllable. Recall also that sac2 has modal voice, the distinctive feature in sac2 is the rising contour, which is identified though the rate of vibration only. Nang2 has some breathiness, which is not very clear in the spectrogram due to the shortness of the tone. Figure 77 shows the pitch graph of sac2 with [ta:k] vs [tak].

Figure 77. Sac2 in [ta:k] and [tak] from Speaker An

![Pitch Graph of Sac2](image)

In Figure 77 we see that the syllable with a short vowel is very short. The tone rises immediately and finishes at approximately 70 ms. The syllable with a long vowel is longer, rising after the first 60 ms. Because the tone is so short and, unlike in sonorant-
final syllables where the nasal has some voicing, there is no voicing here after the oral closure for place of articulation (i.e., -p, -t, -c, -k, -kp), and the tonal feature has to be realized before that point. Figure 78 shows the spectrogram corresponding to Figure 77. Although the rising part cannot be seen in the spectrogram, we still see that the tone has modal voice (regular cycles) and that the short vowel makes the rhyme shorter.

Figure 78. Spectrogram of sac2 in [ta:k] and [tak] from Figure 77

We see the same situation with the tone nang2. Figure 79 shows nang2 in [ta:k] and [tak] from Speaker An. The tone is shorter in the syllable with a short vowel.
In summary, I have shown that tonal features are realized over the whole rhyme, not just on the vowel alone. This is especially clear in non-stop final syllables ending in V:C, the test case required to differentiate between nucleus and rhyme. In stop-final syllables, the tone is too short for its feature to adjust according to vowel length. However, I have argued that it actually does share its tonal feature with the final consonant, namely the feature [obstruent], see Chapter Four. In the following section, one more piece of evidence that the nucleus is not the domain of tone comes from the fact that rhymes tend to have equal lengths.
6.2.3.2. Length in Vietnamese rhyme

Although it is claimed in the literature that the rhyme has a relatively constant length in Vietnamese regardless of its make-up, we saw in the previous section that the equal length is not seen with stop-final consonants (Figures 78 and 79). I focus only on sonorant-final syllables in this section.

Recall that short vowels occur only in closed syllables. In an acoustic study carried out by Gordina and Bystrov 1970 (cited from Hoang Thi Chau 1989:153), the authors claim that the Vietnamese rhyme has inherent length. If the vowel is long, the final consonant is shorter (92a). If the vowel is short, the final consonant is longer (92b).

(92) Length in Vietnamese rhyme

\[
\begin{array}{|c|c|}
\hline
\text{a: } & \text{m} \\
\hline
\text{a} & \text{m} \\
\hline
\end{array}
\]

In the traditional literature, inherent length is used to argue for an intimate relationship between a vowel and a final consonant. Following similar reasoning, it can also be used to argue that the nucleus is not the domain of tone.

My experimental study shows a similar picture, i.e., that the length of syllables with short and long vowels and the same final consonant are usually equal. I use here syllables with the initial [t]. The figures show the spectrogram after the burst of [t], i.e., the vowel and the final sonorant. This part, therefore, is equivalent to the rhyme.

I find a final nasal is long in a syllable with a short vowel. Figures 80 and 81 show the tones ngang and huyen, respectively, in syllables with long and short vowels from
Speaker Phuong. The waveforms of the two syllables with ngang and huyen are shown in the first two rows, respectively. The corresponding spectrograms of the two syllables are in the last two rows.

Figure 80. Same length in CV:N [taːŋ] and CVN [taŋ] with tone ngang

In Figure 80 the final nasal part is represented with very little voicing and weak energy in the spectrograms. The long vowel ends at around 190 ms; the short vowel ends much earlier, at approximately 70 ms. However, the syllable with a long vowel ends at approximately 270 ms, only about 20 ms longer than the syllable with a short vowel, which ends around 250 ms.
In Figure 81, the long vowel ends at approximately 160 ms while the short vowel ends around 90 ms. Both syllables, however, end at almost the same time, at approximately 260 ms. Thus, the final nasal part in these figures, represented with very little voicing, is short if the vowel is long and very long if the vowel is short, i.e., the final sonorant consonant is lengthened to compensate shortness of the short vowel. This is also observed from Speaker Son and Speaker An given earlier in Figures 72 and 76. In these two figures, the syllable with a short vowel has a very long final nasal and its length is almost equal to that of the rhyme with a long vowel. Although it is consistent from the
figures discussed above, that the final nasal is long if the vowel is short, we need to do a larger measurement study to be completely sure about the claim.

What should be noted is that the equal length of rhymes with long and short vowels is a tendency only. Sometimes we find that syllables with short vowels are shorter than syllables with long vowels, but these cases are rather rare. With stop-final syllables, a short vowel makes the rhyme shorter than a long vowel, as seen in Figures 78 and 79. The final voiceless stop is just silence without any voicing, therefore the syllable length relies on the vowel.

In general, syllables, at least sonorant-final syllables, tend to be equal in length. The tendency for a long final sonorant in the syllable if the vowel is short shows that members of the rhyme do 'observe' each other and adjust themselves to preserve a certain length in different syllable types.

In summary, I have shown that tonal features are realized over the whole rhyme, not only on the vowel. This is clearly seen in the syllables with a final sonorant consonant. Moreover, Vietnamese syllables tend to be of a constant length regardless of vowel length. These facts support the claim that the nucleus is not the domain of tone. At the same time, these facts also suggest that the rhyme is the domain because of the way tonal features are distributed in the rhyme and equally realized as the same amount of time in the rhyme. However, we also see evidence from phonological processes, i.e., feature sharing and neutralization that occur in the nucleus only and not in the whole rhyme. Thus they provide evidence for the nucleus - feature sharing - and for the rhyme - the domain of tone and the constituent of equal length.
6.3. Summary

In this chapter I reviewed the literature on the domain of tone and gave evidence that the domain of tone is smaller than the syllable, but bigger than the nucleus. It is smaller than the syllable as a whole because there is no evidence that the initial onset plays a role in tonal realization. It is not a mora, because both long and short vowels are distributed in all syllable types. It is larger than the nucleus, which is made up either of a long vowel or a short vowel with a final. The tonal feature is distributed over the whole rhyme.
Chapter 7

SUMMARY AND IMPLICATIONS

In this chapter I first summarize the major findings of my thesis, and then discuss some predictions and implications of the proposed model. Finally I address questions for future research.

7.1. Summary of major findings

The major contributions of my study on Vietnamese tones are two-fold. First, in Chapter Two I argued for a particular phonological model of tonal representation based on the assumption that markedness is reflected structurally. This model accounts for the phonological patterning of Vietnamese tones, especially in its predictions about neutralization. Second, in Chapter Three I argued based on phonetic evidence that the well-established view of Vietnamese as a pitch register language must be abandoned; what has been called pitch is really the laryngeal features of breathiness and creakiness. This account follows from the phonetic patterning of the tones and, in addition, allows for a simplification of the tonal phonology in reduplication in Vietnamese; the ad hoc and unnatural flip flop rule that was required by previous models is no longer necessary. Importantly, in my model, the phonology of tone is well-grounded in the phonetics.
In addition to these two major findings, I make other contributions that deepen our understanding of Vietnamese tones. First, I present a phonological model of tonal contour that also meshes well with the phonetic facts. Second, I show that the domain of tone is the syllable minus its onset. Finally, I demonstrate that the tonal linear portion is the endpoint, except in curved tones where the middle point is crucial.

7.2. Predictions and implications

7.2.1. Other dialects of Vietnamese

I now turn to some of the predictions that the model makes for other dialects of Vietnamese.

First, the proposed model predicts certain types of neutralization. Not all are attested in the dialect studied, but many more are found when a greater range of dialects is taken into account. One of the predictions of the model is that more complex tones can be neutralized to less complex ones, but less marked tones are not predicted to neutralize to more marked ones. In addition, two types of neutralization are expected, neutralization of contours and neutralization of laryngeal features. These predictions are borne out.

Recall from Chapter Two that the presence of laryngeal features in a tone indicates a more complex tone. When neutralization occurs, these features should be likely candidates for deletion, and this is what is found. For instance, the laryngeal feature of creakiness is not found in Southern dialects. In fact, the tones with this feature in Northern Vietnamese, namely \( n\)ga and \( n\)ang1, neutralize to \( h\)oi and a breathy curved
*nang*1 in Southern Vietnamese, and the Northern creaky *nang*1 is neutralized to a breathy *nang*1 in Central dialects.

In his experimental study, Vu 1982 examines the acoustics of Northern, Central and Southern dialects of Vietnamese. Southern speakers in his study come from Danang to the southern tip of Vietnam. Central speakers come from Hue to Vinh in North Central Vietnam (see map in Appendix 1). Recall that the laryngeal features [constricted] and [spread] are argued to be distinctive in Northern dialects (Chapter Five). This point is important in looking at the phonetics of tones in Central and Southern dialects. Recall that in Northern dialects, the feature [constricted] is realized either as glottal stop or creakiness, and the feature [spread] is realized as breathiness.

Vu summarizes the tonal system of the three major dialects in Table 7.

Table 7. Tones in Vietnamese (Vu 1982)

<table>
<thead>
<tr>
<th>VN names</th>
<th><em>ngang</em>1</th>
<th><em>huyen</em></th>
<th><em>sac</em></th>
<th><em>sac</em> (tac)</th>
<th><em>nang</em>1</th>
<th><em>nang</em> (tac)</th>
<th><em>hoin</em></th>
<th><em>ngai</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>NV</td>
<td>[33]</td>
<td>[21]</td>
<td>[35]</td>
<td>[45]</td>
<td>[21]</td>
<td>[21]</td>
<td>[212]</td>
<td>[325]</td>
</tr>
<tr>
<td>CV</td>
<td>[55]</td>
<td>[42]</td>
<td>[24]</td>
<td>[34]</td>
<td>[31]</td>
<td>[31]</td>
<td>[312]</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>[33]</td>
<td>[21]</td>
<td>[35]</td>
<td>[35]</td>
<td>[212]</td>
<td>[212]</td>
<td>[214]</td>
<td></td>
</tr>
</tbody>
</table>

In this table, 'sac' and 'nang' are *sac*1 and *nang*1, respectively, in my study. Similarly, 'sac tac' (checked *sac*) and 'nang tac' (checked *nang*) are *sac*2 and *nang*2 in my study. NV stands for Northern Vietnam, CV for Central Vietnam, SV for Southern Vietnam. For phonetic notation, Vu uses numeric representations for the relative pitch values and
contour, e.g., ‘5’ is highest and 1 is lowest, 33 is mid level and 21 is low falling. The tones that require attention are those with laryngealization (a term used by Vu to refer to creakiness or glottal closure). These tones, nang and nga in NV and hoi in CV, are underlined.

In Northern dialects, creakiness occurs in nang, [21], and nga, [325], the underlined tones. This is consistent with what I found in my study (see Chapter Three). Now it is important to look at these two tones in other dialects. The tone nga is absent in both Central and Southern dialects. However, unlike in the Northern dialects, the tone nang in Central and Southern dialects does not have creakiness. In Central dialects, creakiness occurs in hoi instead. In Vu’s study, evidence for laryngealization (Vu’s term for creakiness) is shown in this table and in pitch graphs, where laryngealization is represented with dotted lines. Hoang Cao Cuong 1989 also studies tones in different dialects, but he does not describe laryngealization in Central and Southern dialects, focusing only on the pitch range differences between dialects.

In summary, neutralization of the laryngeal feature is found with tones in the sonorant-final syllable type in many dialects of Vietnamese. Nga is neutralized with hoi in these dialects. The tone nang, the more marked member of its pair, does not have creakiness in Central and Southern dialects. In Central dialects, creakiness only appears in hoi, the most complex tone in these dialects. More acoustic work is required on phonation types in these dialects to better understand their tonal systems.

24 In Vu, hoi keeps the feature [curve] as can be seen by the fact that it is represented by three numbers. However, in other studies (e.g., Nguyen and Edmondson 1997, Hoang C 1989), this tone is not curved, but falling. Hoang Cao Cuong 1989 investigates the acoustics of tones in different dialects of Vietnamese. In his study, hoi in the Hue and Do luong dialects, both Central dialects, is a falling tone without the rising portion. Unfortunately, Hoang does not provide any details about the laryngeal features of creakiness and
7.2.2. Language acquisition

I now turn to a second source of evidence for the tonal structures that I propose, namely language acquisition. The proposed model predicts certain patterns of simplification in acquisition, assuming that the acquisition of simpler, or less marked, structures precedes that of more marked ones. Here I consider an example involving the contour as well as the laryngeal aspects of tones. For instance, the proposed structure would predict that a child would acquire the tone nga last, because it is the most complex tone among those in non-stop final syllables. This prediction is borne out. Doan (1977: 111, 113) reports that in his experiment in 1973 with children from two to three years of age, nga is produced as sac1, and hoi is produced like nang1. He concludes that children tend to simplify tones with complex contours to those with simpler contours. Hoang Thi Chau (1989:210) makes a similar observation. She notices that in the early stages of acquisition, nga does not have the curve in production; the resulting tone is something like sac1.

The fact that the contour feature of both hoi and nga is simplified at the early stages of acquisition supports the point that the contour feature [curve], i.e., 'hh', is marked. The fact that children produce nga like sac1 supports the claim that nga, with the register feature [constricted], is more marked than sac1 without the register feature.

---

breathiness in these dialects. It might be that hoi, by losing its complex feature [curve], acquires creakiness in compensation. However, without clear acoustic evidence, this is only speculation.

25 I do not discuss sac2 and nang2 because there is no evidence available on their acquisition. I would expect that the stop-final syllables (which means tones sac2 and nang2) might be learned at later stage for some children.

287
In summary, evidence from the neutralization of hoï and nga in acquisition gives support for the proposed structures of the tones in both register and contour features.  

7.2.3. Neutralization in other languages

In Chapter Two, section 2.2, contour and register nodes are argued to have a sisterhood relationship. Evidence came from reduplicative processes where either node is affected to the exclusion of the other.

Bao 1999a,b provides evidence from Chinese tone sandhi for the separation of contour and register. Interestingly, there is a phenomenon in Beijing Mandarin that can be treated as neutralization of the register feature alone. There are eight tones in Mandarin (Bao 1999b). These are classified into two registers. The classification is shown in (93), from Bao (1999b:10). The two registers are called yin and yang, the four categories that refer to tonal shape are ping (even), shang (rising), qu (departing) and ru (entering). The ru tones occur in stop-final syllables. There are no terms referring to tonal height. The terms ‘departing’ and ‘entering’ are not explained.

---

26 In second language acquisition, phonation seems to give a better cue for distinguishing ngaï from huyen than does tonal height. For instance, in my elementary Vietnamese language class, in order to produce the tone huyen, a breathy level tone, when breathiness is given as a cue, the tone is more likely to be correctly produced than if the student tries to produce huyen as a low tone. Also, in stop-final syllables, the model predicts that sac2, a rising tone, is easier to produce than nang2, a breathy tone: in fact, nang2 is often produced as sac2, but never vice versa, e.g., ngot ‘sweet’, Viet ‘Vietnam’, mat ‘face’. This error is found even in utterance-final position where English speakers do not have a rising intonation in statements.
Traditional classification of tones in Mandarin

<table>
<thead>
<tr>
<th></th>
<th>Even</th>
<th>Rising</th>
<th>Departing</th>
<th>Entering</th>
</tr>
</thead>
<tbody>
<tr>
<td>YIN</td>
<td>ping</td>
<td>shang</td>
<td>qu</td>
<td>ru</td>
</tr>
<tr>
<td>YANG</td>
<td>ping</td>
<td>shang</td>
<td>qu</td>
<td>ru</td>
</tr>
</tbody>
</table>

The phonetics of the registers and tonal shapes (contours) is not provided in the traditional Chinese literature. However, Bao notes that the registers 'yin' and 'yang' serve as a device to capture the phonological patterns of tone sandhi.

In the Beijing Mandarin dialect, the tonal inventory is as in (94) with pitch values indicated. Our interest here is in the organization of the two registers, and therefore the phonetics of tonal shape is not important.

Tonal inventory in Beijing Mandarin

<table>
<thead>
<tr>
<th></th>
<th>Even</th>
<th>Rising</th>
<th>Departing</th>
</tr>
</thead>
<tbody>
<tr>
<td>YIN</td>
<td>ping</td>
<td>55</td>
<td>214</td>
</tr>
<tr>
<td>YANG</td>
<td>ping</td>
<td>35</td>
<td>214</td>
</tr>
</tbody>
</table>

In Beijing Mandarin, there are no final obstruents, and therefore the 'ru' tones do not exist. There are four tones in the inventory. Interestingly, only the two even tones retain a register contrast. The rising and departing tones also remain; however, there is no register contrast in these tones. If, as Bao assumes, the Beijing Mandarin tone system developed from the classical Chinese system, we can see that neutralization of the register feature occurs with these more complex tones without neutralization of tonal contour. The example in (94) shows that there is neutralization between registers and it occurs in the more complex tones (rising and departing) in the system.
The Tianjin Mandarin dialect also shows a similar pattern. The inventory of tones in Tianjin Mandarin is given in (95).

(95) Tonal inventory in Tianjin Mandarin dialect (Bao 1999b)

<table>
<thead>
<tr>
<th></th>
<th>Even</th>
<th>Rising</th>
<th>Departing</th>
</tr>
</thead>
<tbody>
<tr>
<td>YIN</td>
<td>ping 11</td>
<td>shang 24</td>
<td>qu 42</td>
</tr>
<tr>
<td>YANG</td>
<td>ping 55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are only four tones in Tianjin. As in Beijing Mandarin, no register contrast exists outside the even register. In the yang register, there is no shang tone. Bao notes that in this dialect, yang shang is merged into other tonal categories. Only one qu tone is found. If we ignore the phonetic difference between even tones in Beijing and Tianjin (yin ping is high in Beijing but low in Tianjin), we see that the register distinction is lost in qu (departing) tones and in shang (rising) tones. Beijing and Tianjin differ in the placement of the remaining tones. It would be interesting to know why the yang register, but not the yin register, is lost in the rising category. It would also be interesting to know which category the yang shang (rising) tone is merged into. Nevertheless, the major point remains: register neutralization is found without contour neutralization.

7.2.4. Tonal distribution in other languages

Now in terms of tonal distribution, let us see what is expected with respect to markedness. The model proposed in Chapter Two suggests two things: (i) The distribution of tones in stop-final syllables should be more limited than that in nonstop-final syllables, and (ii) if a language has non-level tone(s), it also has level tone(s). Note
that for the languages discussed in this chapter, other than tonal distribution, I do not attempt to give an analysis of tonal patterning because information on the phonological patterning of tones in these languages is not available.

A survey of tonal inventories and of the distribution of tones in other Asian languages shows that both predictions above hold. First, in the tonal inventories of these languages, the number of tones in stop-final syllables, represented as CVT, is smaller than the number in open or sonorant-final syllables, represented as CV(N). Second, although phonetic descriptions are not always given, the labels suggest that non-level tones imply level tones in the system. Ignoring tonal height, I assume that the labels represent the tonal contour, or the tonal shape.

Be (Ong-Be), a Be-Tai language in the Tai-Kadai family (Tran 1999:143), spoken in East and Southeast Asia, has four tones in nonstop-final syllables, but only two tones, high-level and mid-level, occur in stop-final syllables.

(96) Tonal distribution in Be (Hashimoto 1985)

<table>
<thead>
<tr>
<th>BE</th>
<th>mid-level</th>
<th>high-level</th>
<th>low-falling</th>
<th>low-rising</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV(N)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CVT</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sgaw and Pwo are two Karen languages of the Sino-Tibetan family, spoken in the Irrawaddy Delta area and in coastal districts in Lower Burma, the Shan states and Thailand (Benedict 1972). In Sgaw (97) and Pwo (98), there are four tones in non-stop final syllables, but only two tones occur in stop-final syllables.
(97) Tonal distribution in Sgaw (Benedict 1972)

<table>
<thead>
<tr>
<th>SGAW</th>
<th>mid-level</th>
<th>high-level</th>
<th>falling</th>
<th>low-falling</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV(N)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CVT</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(98) Tonal distribution in Pwo (Benedict 1972)

<table>
<thead>
<tr>
<th>PWO</th>
<th>mid-level</th>
<th>low-level</th>
<th>falling</th>
<th>rising</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV(N)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>CVT</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Utsat, a language spoken by a Muslim community in the villages of Yanglan and Huixin near the Southern tip of Hainan, has five tones which occur in non-stop final syllables, but only four tones in obstruent-final syllables (Maddieson and Pang 1993:75).

(99) Tonal distribution in Utsat (Maddieson & Pang 1993)

<table>
<thead>
<tr>
<th>UTSAT</th>
<th>high</th>
<th>mid</th>
<th>low</th>
<th>falling</th>
<th>rising</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV(N)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CVT</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Finally, Yay, a Northern Tai language spoken in North Vietnam, has six tones. All can occur in non-stop final syllables. In non-stop final syllables, the high rising-falling and high-rising tones occur very infrequently (Gedney 1965:181).
From these inventories, we see that, in general, tone has a wider distribution in non-stop final syllables, and the presence of simpler contours implies that of more complex contours.

7.3. Questions for future research

In the final section, I would like to suggest some problems and questions that arise from my analysis of Vietnamese tones.

7.3.1. A mismatch in Chaozhou tones

I begin by examining a Chinese language which has a mismatch problem similar to that discussed in Chapter Two.

Chaozhou is a Southern Min dialect spoken in Fujian and Guangdong provinces and Southeast Asia (Bao 1999a). Bao introduces tone sandhi facts in this dialect to argue for the separation of tonal register and contour. There are six phonemic tones in this dialect. The specification of the six tones according to Bao is as in (101). All tones except 33 and 11 undergo tone sandhi in non-phrase-final position. The tone 35 is left unspecified because it is problematic: it does not behave as expected in tone sandhi.

---

(100) Tonal distribution in Yay (Gedney 1965)

<table>
<thead>
<tr>
<th></th>
<th>level</th>
<th>low-level</th>
<th>mid-rising</th>
<th>falling</th>
<th>high-rising</th>
<th>high rising-falling</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV(N)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CVT</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

293
(101) Specification of Chaozhou tones

a. 33  L, h     d. 55  H, h
b. 53  H, hl    e. 35

c. 213 L, lh    f. 11  L, l

L, H represent register and h, l represent contour. Tone sandhi is regarded as Register harmony in which ‘the pitch height of the sandhi tone - its register - is determined by the register of the following tone’ (Bao 1999a: 487). I will discuss the mismatch problem of tone 35 without going into details of sandhi and Bao’s analysis. The patterns of tones 53 and 213 are given in (102). In (102a), the low register tone 213 becomes a high register tone 53 before a high register tone 53 or 55, or it becomes 42 before a low register tone.27

(102) Tone sandhi in Chaozhou

a. 213 ----> 53/33, 55
   42/33, 213, 11, 35

b. 53 ----> 35/33, 55
   24/33, 213, 11, 35

In (102b), the high register tone 53 remains a high register tone (although it changes from

27 It is not clear how the register of the surface tone 42 is determined, and what determines a register is ‘high’ or ‘low’. Note that tones in the environment in (113a) include 35, a high tone.
a falling contour 53 to a rising contour 35) before a high tone. It becomes a low register tone 24 before the low register tones 33, 213, 11 and 35. The tone 35 is important here. According to Bao, 'on the surface, it is a high rise, which should be specified as H, Ih, but it triggers pitch lowering' (1999a: 490). Therefore, 35 causes a mismatch. Phonetically it is a high tone, but phonologically it patterns as if it were a low tone. This problem seems similar to the problem we have seen in Vietnamese: hoi is claimed to be a low tone but it behaves as if it were a high tone. However, we saw in Chapter Three that hoi is in fact not lower than nga, its counterpart.

Bao suggests a solution for the problematic behaviour of tone 35 by assuming that the citation tones are not underlying (35 is a citation tone), but the sandhi tone is more basic. Therefore, 35 should be treated as being derived from the underlying 24 after the application of Register harmony. This solution runs into another difficulty (see Bao 1999a for details). A solution to the entire problem might be found by investigating the acoustics of Chaozhou tones. There might be something in the tone 35 that makes it phonetically a low tone.

This suggestion raises a few more interesting questions concerning phonation types. In the following section I address some typological questions concerning Vietnamese dialects, and then other tone languages.

7.3.2. What are the underlying features in all dialects of Vietnamese? Implications for perception and typology

In this study I have examined the tonal properties of Northern Vietnamese. Since the tonal inventories differ between dialects (see Table 7, section 7.2.1), an obvious line
of research involves the study of tone in these other dialects. In particular, which features, if any, are stable across all dialects? All researchers agree that it is not tonal height. Vu 1982 suggests that it is tonal shape. However, a close look at Hoang’s pitch graphs reveals that what is called the same tone has very different shapes in different dialects (see Chapter Four, section 4.1.1). Hoang Thi Chau (1989:215) observes that in all dialects, the ‘low’ tones are very close together, occurring as a bundle. What helps a listener distinguish each tone in a bundle of tones? She suggests that it is the contrastive relationship between tones in terms of register and contour. This relationship is shown in the phonology of the tonal system represented by tonal patterning in reduplication. Hoang Thi Chau’s grounding of the perception of tones in terms of contrast seems correct, but she does not offer an explanation in terms of just what these contrastive features are. Tonal height is questionable, given the often marginal differences between tones.

In Hoang Cao Cuong 1989, tonal height is compared between dialects. He notices that the pitch-span of tones, i.e., the maximal and minimal values in a speaker’s pitch range, differ across dialects. For instance, in Hue and Do Luong (Central dialects), this range is narrower than in North and South Vietnamese dialects. Generally in North and South Vietnamese, tones are close to each other at the beginning. The difference in F0 between tones is very large only towards the end. In Central dialects, tones do not vary much in their F0, even towards the end point. According to Hoang, this leads to an impression by speakers from other dialects that Central tones sound ‘monotone’ and ‘heavy’ because the F0 does not change much (1989:16). The narrow pitch range is often regarded as being responsible for confusions by speakers from other dialects. Hoang
suggests that perhaps the development of Vietnamese tones from segmental features is still in progress in the Central dialects; therefore we might find other phonation characteristics of tones, a relic of historical segments, in these dialects. In Central dialects, breathiness is particularly strong. My hypothesis is that breathy voice in Central dialects functions as a trade-off for pitch differences: smaller pitch differences will correlate with greater phonation differences. This follows from the fact that phonation types can be primary features over pitch height, as seen in the Hanoi dialect. It would be interesting to investigate the phonation types of tones in Central Vietnam to test this hypothesis.

The assumption that tonal height is a primary cue in tone languages leads to an interesting question regarding perception. If in the Central dialects, the narrow pitch-span might be responsible for confusions by speakers of other dialects, why does this confusion not happen in the same way with males in Northern dialects? Compared to females, males in Northern dialects have a very narrow pitch-span, especially with the level pair ngang and huyen. In Chapter Three, we saw that the pitch range is quite different between females and males, and even between speakers of the same sex. In some male speakers (e.g., Son), the difference in F0 between tones is very small, e.g., from about 5 to 8 Hz difference between ngang and huyen, as seen in Figure 82, repeated from Chapter Three.
This is typical of ngang and huyen in Speaker Son. If the difference in F0 is significant, is he less intelligible than other men? Or are men in general less intelligible than women? If not, what makes Vietnamese male speech as intelligible as Vietnamese female speech? In Central dialects, the problem is exacerbated: the pitch range is very narrow for both men and women. This raises an interesting question for acoustic and auditory phonetics: How do tones in male speakers in Central dialects differ from those in male speakers in the North? How does a listener adjust his/her own scale to perceive different pitch ranges in a tone language?

My instrumental work has interesting implications for the study of the perception of Vietnamese tone. I expect that although speakers of different dialects have different acoustic characteristics in producing tones, they share the same perception of tones. What features are crucial in perceiving a tone? Can tones be discriminated in the
first 40-60 ms? A perception test with speakers from different dialects would shed light on this issue. Do breathiness and creakiness really provide the information claimed, i.e., are perceptual contrasts present if we smooth them over?

Finally, what types of languages are expected typologically? We predict at least three types: (i) languages that use contour features (including tonal height); (ii) languages that use phonation types; (iii) and languages that use both distinctively. Languages, thus, differ in which parameter they choose to be primary. It is well known that African tone languages use pitch height as the primary, and even the only, feature (e.g., when discussing tonal features in African languages, Odden (1995) mentions only pitch). This tonal typology can also be found in the East Asian language area (Thongkum, 1988:319). Note that the discussion below is based mainly on the phonetics of tone. No phonological behaviour of tones is discussed in the sources.

For the first type, pitch height only, Charoenma (1983) reports that Wiang Papao Lua, also called Khamet or Lawa, a Palaung-wa language of the Mon-Khmer family spoken in Chiangrai, Thailand, has two contrastive tones (tone is equal to pitch height in his terms): falling breathy and normal tone. The breathy tone is mid falling, and the normal tone varies freely between low-rising and mid-level. Voice quality is not contrastive in Lua. For the second type, with only phonation type distinctive, Charoenma reports in the same study that Lamet, another Palaung-wa language spoken in Lampang, Laos, contrasts two voice registers. Unlike Wiang Papao, pitch height is not distinctive in Lamet, but is conditioned by syllable structure. For the third type, Vietnamese suggests that both phonation and contour features are employed in this language. Eastern Tamang, a Tibeto-Burman language, provides another case of this
type. In this language, there are four tones, of which two have clear voice (modal voice) and two with breathy voice (Weidert 1987). There is an interaction between tone, initial voicing and voice quality. However, because the four tones are contrastive in sonorant-initial syllables, they should be treated as distinctive, and details of initial voicing are omitted here. The four tones are shown in (103).

(103) Clear phonation: \( H \) (or \( HL \))

\[ M \]

Breathy phonation: \( LM \)

\[ L \]

Two things are worth noting here: ‘the two breathy tones are lower in pitch than the two clear tones, and within a single phonation type, the tones are distinguished by level or shape’ (Yip 1995:486). First, it is not surprising to see that breathy tones are lower than clear tones because we have seen that breathiness lowers a tone in Chapter Three. Second, if within a single phonation, the tones are distinguished by their shape, it could be that both breathiness and contour are used as distinctive features in this language, and we do not need to refer to tonal height here. It would be interesting to examine both the acoustics and the phonological patterns of tone in Eastern Tamang.

The questions raised in this section are not simple ones, but they would likely give extremely interesting answers. The Vietnamese case suggests that there remains much to understand about tones.
7.4. Conclusion

In pursuing questions that have been raised from the beginning of the thesis, I have argued that Vietnamese tones are organized in a hierarchical structure, that laryngeal features of phonation are distinctive, that pitch height is not distinctive, and that the features are grounded phonetically.

By proposing a structural representation for Vietnamese tones, this study is able to capture systematically the tonal patternings seen in the language. The model can also predict various types of neutralization. Another equally important contribution is that replacing pitch height with phonation types and tonal shape gives a natural and elegant account of reduplication, the most widely discussed evidence for the patterning of tones in the traditional literature. When the assumption that tone is pitch height is abandoned, there emerges a new way to examine tonal languages, contributing to our understanding of just what it means to be tonal.

Finally, if ‘tone’ is equal to ‘pitch’, Vietnamese - perhaps along with many others - is not a ‘tone language’ after all.
REFERENCES


Cao, Xuan Hao 1998. Hai cach mieu ta he thong thanh dieu tieng Viet (Two ways to describe the tonal system in Vietnamese) in Tieng Viet - May van de ngu am, ngu phap, ngu nghia (Some issues on Vietnamese phonology, syntax, semantics) . Collection of articles by Cao Xuan Hao. Hochimin city: Nha xuat ban Giao duc.


Cu, Tu D., Hoang Van Thung & Nguyen Nguyen Tru. 1977. Ngu am hoc tieng Viet hien dai (Phonology of the modern Vietnamese). Hanoi: Publisher of Education.


Han, Mieko 1969. Studies in the phonology of Asian languages VIII: Vietnamese tones. California: University of Southern California


Hoang, Dung. 1998. Ban them ve van de nhan dien tu lay tieng Viet (On the classification of Vietnamese reduplication), ms. Hanoi University (?).


Vuong, Le H. & Hoang Dung. 1994. Ngu am tieng Viet (Vietnamese phonology). Hanoi. DHSP.


APPENDIX 1

Map of Vietnam
### APPENDIX 2

Initial inventory of Vietnamese Northern dialects

<table>
<thead>
<tr>
<th></th>
<th>labial</th>
<th>alveolar</th>
<th>palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop</td>
<td>t</td>
<td>c</td>
<td>k</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>f</td>
<td>s</td>
<td>x</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td></td>
<td>v</td>
<td>z</td>
<td></td>
<td>γ</td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td>ṭ</td>
<td>ṭ</td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td></td>
<td></td>
<td>l</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX 3

Subjects (information at the time of recording, during the summer 2000)

1. Nguyen Thi Van Khanh, female, 50 years old. Born in Ha Dong, close to Hanoi. Father came from Ha Dong, mother from Ha Noi. Left for Saigon in childhood and lived in a northern community. Came to Canada in 1981. Teach Vietnamese to elementary school students of the heritage program, Toronto Board of Education.

2. Nguyen Van Hoang, male, 23 years old. Born and grew up in Hanoi, third year student at University of Toronto in 2000. He came to Canada in 1996.


7. Nguyen Vinh Hung, male, 40 years old. Born and grew up in Hanoi. Came to Canada in 1990. Received a graduate degree at University of Toronto. Pharmacist.

8. Nguyen Thi Thanh Binh, female, 42 years old. Born and grew up in Hanoi. Moved to live in Ha Nam Ninh from 6 to 17, then back to Hanoi from 1980 until now. A visa graduate student at University of Toronto.

APPENDIX 4
Sentence list

1. Chữ diễm phải ngay.
   Chữ cảo phải ngay.
   Chữ tây phải ngay.
   Chữ tài phải ngay.
   Chữ cả phải ngay.
   Chữ mái phải ngay.
   Chữ tăng phải ngay.
   Chữ täy phải ngay.
   Chữ xã phải ngay.
   Chữ tiếu phải ngay.

2. Chữ tấm phải ngay.
   Chữ tăng phải ngay.
   Chữ mà phải ngay.
   Chữ đặc phải ngay.
   Chữ tang phải ngay.
   Chữ cả phải ngay.
   Chữ dào phải ngay.
   Chữ tài phải ngay.
   Chữ các phải ngay.
   Chữ điện phải ngay.

3. Chữ tấm phải ngay.
Chữ xác phải ngay.
Chữ ma phải ngay.
Chữ cá phải ngay.
Chữ tái phải ngay.
Chữ tây phải ngay.
Chữ tái phải ngay.
Chữ mái phải ngay.
Chữ xá phải ngay.
Chữ tiếng phải ngay.

4. Chữ tanı phải ngay.
Chữ tài phải ngay.
Chữ ca phải ngay.
Chữ tăng phải ngay.
Chữ cái phải ngay.
Chữ tải phải ngay.
Chữ tang phải ngay.
Chữ dài phải ngay.
Chữ dã phải ngay.
Chữ điểm phải ngay.

5. Chữ điều phải ngay.
Chữ tăng phải ngay.
Chữ cái phải ngay.
Chữ mao phải ngay.
Chữ mạc phải ngay.
Chữ xải phải ngay.
Chữ tay phải ngay.
Chữ ta phải ngay.
Chữ tảo phải ngay.
Chữ tiếng phải ngay.

6. Chữ tán phải ngay.
Chữ táy phải ngay.
Chữ mả phải ngay.
Chữ ta phải ngay.
Chữ tảng phải ngay.
Chữ tăng phải ngay.
Chữ xảo phải ngay.
Chữ tá phải ngay.
Chữ táng phải ngay.
Chữ điện phải ngay.

7. Chữ tằm phải ngay.
Chữ tai phải ngay.
Chữ da phải ngay.
Chữ dãi phải ngay.
Chữ dá phải ngay.
Chữ mả phải ngay.
Chữ xà phải ngay.
Chữ  mão  phải ngay.
Chữ  tằng  phải ngay.
Chữ  diếng  phải ngay.

8.  Chữ  tiến  phải ngay.
Chữ  tác  phải ngay.
Chữ  táo  phải ngay.
Chữ  cạc  phải ngay.
Chữ  mạc  phải ngay.
Chữ  tà  phải ngay.
Chữ  dạc  phải ngay.
Chữ  tằng  phải ngay.
Chữ  tà  phải ngay.
Chữ  tán  phải ngay.

Chữ  tiêu  phải ngay.
Chữ  xa  phải ngay.
Chữ  dạo  phải ngay.
Chữ  tác  phải ngay.
Chữ  xác  phải ngay.
Chữ  xạ  phải ngay.
Chữ  xác  phải ngay.
Chữ  đạ  phải ngay.
Chữ  tiêu  phải ngay.
10. Chủ tần phải ngày.
    Chủ thang phải ngày.
    Chủ đặc phải ngày.
    Chủ cà phải ngày.
    Chủ xãi phải ngày.
    Chủ xáo phải ngày.
    Chủ tác phải ngày.
    Chủ dà phải ngày.
    Chủ má phải ngày.
    Chủ tiêm phải ngày.

11. Chủ điện phải ngày.
    Chủ tà phải ngày.
    Chủ dạ phải ngày.
    Chủ xác phải ngày.
    Chủ xã phải ngày.
    Chủ cạo phải ngày.
    Chủ tác phải ngày.
    Chủ đặc phải ngày.
    Chủ cả phải ngày.
    Chủ tần phải ngày.
    Chủ mặc phải ngày.
    Chủ điển phải ngày.
CV - CV:T, CVT

\[
\begin{array}{cccccccc}
\text{ngang} & \text{sac}l & \text{huyen} & \text{nang}l & \text{hoi} & \text{nga} & \text{sac}2 & \text{nang}2 \text{sac}2 & \text{nang}2 \\
\text{ma} & \text{má} & \text{mà} & \text{mã} & \text{mà} & \text{mã} & \text{má} & \text{mà} & \text{mã} \\
\text{ta} & \text{tá} & \text{tà} & \text{tã} & \text{tà} & \text{tã} & \text{tác} & \text{tác} & \text{tác} \\
\text{xa} & \text{xá} & \text{xà} & \text{xã} & \text{xà} & \text{xã} & \text{xác} & \text{xác} & \text{xác} \\
\text{da} & \text{tá} & \text{tà} & \text{dã} & \text{dã} & \text{dã} & \text{dác} & \text{dác} & \text{dác} \\
\text{ca} & \text{cá} & \text{cà} & \text{cã} & \text{cã} & \text{cã} & \text{các} & \text{các} & \text{các} \\
\end{array}
\]

Long [V:] and short [V] vowels

\[
\begin{array}{cccccc}
\text{ngang} & \text{sac}l & \text{huyen} & \text{nang}l & \text{hoi} & \text{nga} \\
\text{[V:] tang} & \text{tang} & \text{tàng} & \text{tàng} & \text{tàng} & \text{tàng} \\
\text{[V] tàng} & \text{tàng} & \text{tàng} & \text{tàng} & \text{tàng} & \text{tàng} \\
\text{[V:] tai} & \text{tái} & \text{tài} & \text{tài} & \text{tài} & \text{tài} \\
\text{[V] tay} & \text{táy} & \text{tày} & \text{tày} & \text{tày} & \text{tày} \\
\end{array}
\]

CV:G - CV:N/ CVN (G = /j/, /w/)

\[
\begin{array}{cccc}
\text{hoi} & \text{nga} & \text{hoi} & \text{nga} \\
\text{[V:] mái} & \text{má} & \text{máo} & \text{máo} \\
\text{tái} & \text{tái} & \text{tào} & \text{tào} \\
\text{xái} & \text{xái} & \text{xào} & \text{xào} \\
\text{cái} & \text{cái} & \text{cào} & \text{cào} \\
\text{dái} & \text{dái} & \text{dào} & \text{dào} \\
\text{[V:] tâm} & \text{tâm} & \text{tán} & \text{tán} \\
\text{[V] tâm} & \text{tâm} & \text{tán} & \text{tán} \\
\end{array}
\]

316
CVVG - CVVN  (VV: diphthong)

hoi nga hoi nga hoi nga hoi nga

tiểu tiém tién tiếng tiếng

diều diễm điện điện diếng diếng
APPENDIX 5

Pitch graphs of others

1/ Speaker Phuong (female): six tones with [ka], and sac2 and nang2 with [ka:k]
2/ Speaker Khanh (female): six tones with [ka], and sac2 and nang2 with [ka:k]
3/ Speaker Dung (female): six tones with [ka], and sac2 and nang2 with [ka:k]
4/ Speaker Van (female): six tones with [ka], and sac2 and nang2 with [ka:k]
5/ Speaker An (female): six tones with [ka], and sac2 and nang2 with [ka:k]
Speaker Hung (male): six tones with [ka], and sac2 and nang2 with [ka:k]
Speaker Hongo (male): six tones with [ka], and 442 and 142 with [ka:k]
8/ Speaker Son (male): six tones with [ka], and sac2 and nang2 [ka:k]
APPENDIX 6

Other speakers: spectrograms of six tones with [ka] and sac2 and nang2 with [ka:k]

1/ Speaker Phuong: ngang-huyen, sac1-nang1
Phuong: hoi-nga, sac2-nang2
2/ Speaker Khanh: ngang-huyen, sac1-nang1
Khanh: hoi-nga, sac2-nang2
3/ Speaker Dung: ngang-huyen, sacl-nang1
Dung: hoi-nga, sac2-nang2
4/ Speaker Van: ngang-huyen, sac1-nang1
Van: hoi-nga, sac2-nang2
Speaker An: ngang-huyen, sac1-nang1
An: hoi-nga, sac2-nang2
6/ Speaker Hung: ngang-huyen, sac1-nang1
Hung: hoi-nga, sac2-nang2
7/ Speaker Hoang: ngang-huyen, sac1-nang1
Hoang: hoi-nga, sac2-nang2
8/ Speaker Binh: ngang-huyen, sac1-nang1
Binh: hoi-nga, sac2-nang2